

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 09:27:12 ; Search time 1854.65 Seconds

(without alignments)
16507.481 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463
1 caagagaataagctcgcgcc.....aaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GeneBml:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
--------	-------	-------	-------	--------	----	----	-------------

1	1463	100.0	1463	6	ARI52403		
2	1421.8	97.2	1478	6	ARI52406		ARI52403 Sequence
3	1409.8	96.4	1541	6	ARI52404		ARI52406 Sequence
4	1350.2	92.3	1381	6	ARI52405		ARI52404 Sequence
5	927.4	63.4	1354	6	AR021666		ARI52405 Sequence
6	326.8	22.3	1469	5	XL068141		AR021666 Oriza sat
7	308.6	21.1	1461	5	XL064563		U68141 Xenopus lae
8	307	21.0	1357	5	AF065397		U64563 Xenopus lae
9	305.4	20.9	1149	5	AF036327		AF065397 Xenopus 1
10	305.4	20.9	1149	5	AF036327		AF036327 Xenopus 1
11	305.4	20.9	1149	5	AF036327		BC010203 Mus muscu
12	302.8	20.7	1757	9	AC026761		AC026761 Mus muscu
13	302.8	20.7	2031	9	BC000323		X76771 Homo sapien
14	302.8	20.7	138627	2	AP000591		BC000323 Homo sapi
15	302.8	20.7	169053	2	AP002380		AP000591 Homo sapi
16	302.8	20.7	185035	2	AC004470		AP002380 Homo sapi
17	302.8	20.7	196080	9	AC004228		AC004470 Homo sapi
18	302.6	20.7	1144	5	HUMENB1A		AC004228 Homo sapi
19	299	20.4	1146	5	AB058602		L37374 Homo sapien
20	282.4	20.0	39563	9	HSU73629		AB058602 Gallus ga
21	288.2	19.7	2192	10	AF281018		U73629 Human chrom
22	279.8	19.1	2033	10	MUSFENIX		AF281018 Rattus no
23	278.2	19.0	6361	10	AY014962		L26320 Mouse flap
24	274.6	18.8	2114	8	SCYKL113C		AY014962 Mus muscu
25	274.6	18.8	10660	8	S93804		Z28113 S.cerevisia
26	259.4	17.7	2299	8	SPRAD2		S93804 BAF1/ABF1/Y
27	259.4	17.7	28491	8	SPRAD2		X77041 Saccharomyc
28	259	17.7	151865	2	AL356236		Z29167 S.pombe chr
29	259	17.7	179461	2	AC019177		AL356236 Homo sapi
30	257.4	17.6	173256	2	AC073301		AC019177 Homo sapi
31	257.4	17.6	178806	2	AC013557		AC073301 Homo sapi
32	252.6	17.3	37688	2	AL513222		AC013557 Homo sapi
33	252.6	17.3	37688	2	DMCEG0003		AL513222 Drosophi
34	252.6	17.3	95850	3	AC020063		AL031863 Drosophi
35	252.6	17.3	154985	3	AC004641		AC020063 Drosophi
36	252.6	17.3	160338	3	AC099021		AC004641 Drosophi
37	252.6	17.3	270306	3	AE003805		AC099021 Drosophi
38	241.8	16.5	194322	2	AC024791		AE003805 Drosophi
39	241.8	16.5	319857	2	AC006782		AC024791 Caenorhab
40	237	16.2	2783	3	AF093702		AC006782 Caenorhab
41	237	16.2	234112	3	PFMALP2		AF093702 Plasmodi
42	236.6	16.2	1953	3	AF278764		AF093702 Plasmodi
43	236.4	16.2	194439	3	CNS07BGA		AF278764 Plasmodi
44	233.6	16.0	120848	2	AC007801		AL590443 Chromosom
45	159.6	10.9	1023	6	AR062093		AC007801 Drosophi

ALIGNMENTS

RESULT 1
LOCUS ARI52403 1463 bp DNA
DEFINITION Sequence 1 from patent US 6232527. linear
ACCESSION ARI52403
VERSION ARI52403.1 GI:15118453 PAT 08-AUG-2001
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1463)
AUTHORS Mahajan,P.B.
TITLE Maize Rad2/REN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 1 15-MAY-2001;
FEATURES
source Location/Qualifiers
BASE COUNT 466 a 292 c 361 g 344 t
ORIGIN

Query Match 100.0% Score 1463; DB 6; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CACGAGATAGCTCGCGCGCGGTTCTTGGCCACTCCGGCTCA GCCCGCCGCCCA 60
    |||
    |||
    |||
Db 1 CAGAGAAATAGCTCGCGCGCGGTTCTTGGCCACTCCGGCTCA GCCCGCCGCCCA 60
OY 61 CCCGCCACAGCCCGCCGACGATGGGCTTCAGAGGTTTACGAAATGCTGGGCGGAC 120
    |||
    |||
    |||
Db 61 CCCGCCACAGCCCGCCGACGATGGGCTTCAGAGGTTTACGAAATGCTGGGCGGAC 120
OY 121 AATGGCGCCAAAGCGATGAAGAGAGATTGAGAGCTACTTCGCGCCGCAAAATCGCC 180
    |||
    |||
    |||
Db 121 AATGGCGCCAAAGCGATGAAGAGAGATTGAGAGCTACTTCGCGCCGCAAAATCGCC 180
OY 181 GTGAGCGCCAGCATGAGCATATACCACTTCTATTGTAGTGAAGAGACAGCATGAAA 240
    |||
    |||
    |||
Db 181 GTGAGCGCCAGCATGAGCATATACCACTTCTATTGTAGTGAAGAGACAGCATGAAA 240
OY 241 ACTGTCACAAATGAAGCTGTGGAAGTCACTAGTCATTTGCAAGGAATGTTCAACCGGACA 300
    |||
    |||
    |||
Db 241 ACTGTCACAAATGAAGCTGTGGAAGTCACTAGTCATTTGCAAGGAATGTTCAACCGGACA 300
OY 301 ATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTTATGGCAAGCTCTGAT 360
    |||
    |||
    |||
Db 301 ATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTTATGGCAAGCTCTGAT 360
OY 361 ATGAAGAAACAAGAGCTTGTCTAAAGATCTCAAAAAGAGATGTGCAACCAAGATCTG 420
    |||
    |||
    |||
Db 361 ATGAAGAAACAAGAGCTTGTCTAAAGATCTCAAAAAGAGATGTGCAACCAAGATCTG 420
OY 421 ACTGAGCAGTAGAGTAGAGATGAAGATGCGATTGAAAAATTTAGACCAAGAGACTGTA 480
    |||
    |||
    |||
Db 421 ACTGAGCAGTAGAGTAGAGATGAAGATGCGATTGAAAAATTTAGACCAAGAGACTGTA 480
OY 481 AAGGTCACAGGCAACACACAGGAATTTGAACGCGCTATTAGACTTTAGGGGGTTCCT 540
    |||
    |||
    |||
Db 481 AAGGTCACAGGCAACACACAGGAATTTGAACGCGCTATTAGACTTTAGGGGGTTCCT 540
OY 541 GTTGTAGAGGAGACCTTTCGAGCAGGAAGCAGATGTGACGCCCTTTCATTAAGATTAAG 600
    |||
    |||
    |||
Db 541 GTTGTAGAGGAGACCTTTCGAGCAGGAAGCAGATGTGACGCCCTTTCATTAAGATTAAG 600
OY 601 GTGTGCGTGTGCTTCGAGAGATATGACCTCTTACTTTTGGGGCTCCAGCGTTCCTT 660
    |||
    |||
    |||
Db 601 GTGTGCGTGTGCTTCGAGAGATATGACCTCTTACTTTTGGGGCTCCAGCGTTCCTT 660
OY 661 CGTCAATTAATGATCCCAAGTTCACAGAAATACCTGTGATGGAATTTGATGTCGCAAG 720
    |||
    |||
    |||
Db 661 CGTCAATTAATGATCCCAAGTTCACAGAAATACCTGTGATGGAATTTGATGTCGCAAG 720
OY 721 GTTTGGAGAGACTTGAAGTCCATGAGCAGTTCATTTGATGTCGTCGTCGTCGTCGTCG 780
    |||
    |||
    |||
Db 721 GTTTGGAGAGACTTGAAGTCCATGAGCAGTTCATTTGATGTCGTCGTCGTCGTCGTCG 780
OY 781 TGTGACTATTGTGATGATGATCAAGAGTATCGGGGGGCAAAAGCTCTGAAATCTTATTCGT 840
    |||
    |||
    |||
Db 781 TGTGACTATTGTGATGATGATCAAGAGTATCGGGGGGCAAAAGCTCTGAAATCTTATTCGT 840
OY 841 CAACATGGGTCCATGAAGAGCATCTTGGAGATCTTAATTAAGACAGATATCAAAATTCCT 900
    |||
    |||
    |||
Db 841 CAACATGGGTCCATGAAGAGCATCTTGGAGATCTTAATTAAGACAGATATCAAAATTCCT 900
OY 901 GAGGACTGGCTTTACCAAGAAAGCTGAGCGTTGTTCAAGAGACCTTAATGTCAATTTGAT 960
    |||
    |||
    |||
Db 901 GAGGACTGGCTTTACCAAGAAAGCTGAGCGTTGTTCAAGAGACCTTAATGTCAATTTGAT 960
OY 961 ATTCTGAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
    |||
    |||
    |||
Db 961 ATTCTGAGCTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
OY 1021 GATTAATGTTTCAAGAGATGCGGTGACAAAGGCCATAGAGAATCAAAATCTGCCAAG 1080
    |||
    |||
    |||
Db 1021 GATTAATGTTTCAAGAGATGCGGTGACAAAGGCCATAGAGAATCAAAATCTGCCAAG 1080

```

```

OY 1081 AATAATCGTCGCAAGAGAGACTGAGTCCCTTTTTCGAAGCCAACTGCCACACATGACGA 1140
    |||
    |||
    |||
Db 1081 AATAATCGTCGCAAGAGAGACTGAGTCCCTTTTTCGAAGCCAACTGCCACACATGACGA 1140
OY 1141 CCGCTAAACGGAGAGAGACTTCGGATTAACCAAGCAGGACGCTGGGACCAAGAAAACA 1200
    |||
    |||
    |||
Db 1141 CCGCTAAACGGAGAGAGACTTCGGATTAACCAAGCAGGACGCTGGGACCAAGAAAACA 1200
OY 1201 AAGCTGTGGGAAAGAAATTAATCTTGAATGCTGATGTACAACTACGATCAAGAAAG 1260
    |||
    |||
    |||
Db 1201 AAGCTGTGGGAAAGAAATTAATCTTGAATGCTGATGTACAACTACGATCAAGAAAG 1260
OY 1261 CAGCGTGGCGGTGATCACTTCGCTTAATTTAACTCCCTGTTTTAACTCAGAGCTTT 1320
    |||
    |||
    |||
Db 1261 CAGCGTGGCGGTGATCACTTCGCTTAATTTAACTCCCTGTTTTAACTCAGAGCTTT 1320
OY 1321 GGTAAATTTGCTCANTTTCAAGCTGGGGTAAAGTAAAGTGTGTTTGAAGAGATTGGTG 1380
    |||
    |||
    |||
Db 1321 GGTAAATTTGCTCANTTTCAAGCTGGGGTAAAGTAAAGTGTGTTTGAAGAGATTGGTG 1380
OY 1381 TACCAAGTAAACAACTATGCGTGTGTTTTTAACTTCTTCTTGAAGTAAAAAAA 1440
    |||
    |||
    |||
Db 1381 TACCAAGTAAACAACTATGCGTGTGTTTTTAACTTCTTCTTGAAGTAAAAAAA 1440
OY 1441 AAAAAAAAAAAAAAAAAAAAAA 1463
    |||
    |||
    |||
Db 1441 AAAAAAAAAAAAAAAAAAAAAA 1463

```

RESULT 2
 ARI52406
 LOCUS ARI52406 1478 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 7 from patent US 6232527.
 ACCESSION ARI52406
 VERSION ARI52406.1 GI:15118456
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1478)
 AUTHORS Mahajan, P. B.
 TITLE Maize Rad2/FEN-1 orthologues and uses thereof
 JOURNAL Patent: US 6232527-A 7 15-MAY-2001;
 FEATURES
 source 1..1478
 /organism="unknown"
 BASE COUNT 463 a 302 c 365 g 348 t
 ORIGIN

Query Match 97.2%; Score 1421.8; DB 6; Length 1478;
 Best Local Similarity 98.5%; Pred. No. 0;
 Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

OY 7 AATAGCTCGCGCGCGGTTCTTGGCCACTCCGGCTCA GCCCGCCGCCCA 66
    |||
    |||
    |||
Db 19 AATAGCTCGCGCGGTTCTTGGCCACTCCGGCTCA GCCCGCCGCCCA 78
OY 67 ACAAGCGCCGACAGAGATGGGCTCAAGAGGTTTACGAAACTGTGGCGGACATATGCG 126
    |||
    |||
    |||
Db 79 ACAAGCGCCGACAGAGATGGGCTCAAGAGGTTTACGAAACTGTGGCGGACATATGCG 138
OY 127 CCCAAGCGCATGAAGAGAGCAAGTTGAGAGCTACTGCGCGCAAAATCGCCGTCGAC 186
    |||
    |||
    |||
Db 139 CCCAAGCGCATGAAGAGAGCAAGTTGAGAGCTACTGCGCGCAAAATCGCCGTCGAC 198
OY 187 GCCAGCATGAGCATATACAGTTCCTGATTTGATTTGGAAGAGCAGGCAATGGAATCTTC 246
    |||
    |||
    |||
Db 199 GCCAGCATGAGCATATACAGTTCCTGATTTGATTTGGAAGAGCAGGCAATGGAATCTTC 258
OY 247 ACAATAGAGCTGTGTAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAAATAGA 306
    |||
    |||
    |||
Db 259 ACAATAGAGCTGTGTAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAAATAGA 318

```


Db	613	GAAATATGGAACTCCCTTACTTTTGGGGCTCCACGGTTCCTTGGTCAATTTAATGATACCA	672		
Qy	679	AGTTCCAAAGAAATACCTGTGATGAAATTTGATGTTGCCAAGTCTTTGGAGAGACTTGA	738		
Db	673	AGTTCCAAAGAAATACCTGTGATGAAATTTGATGTTGCCAAGTCTTTGGAGAGACTTGA	732		
Qy	739	CTCACCATGGACGATTCATGATTTGGTATGACATCCGTGTGTGATGTGATCTATTTGTGATAC	798		
Db	733	CTCACCATGGACGATTCATGATTTGGTATGACATCCGTGTGTGATGTGATCTATTTGTGATAC	792		
Qy	799	ATCCAAAGCTTTCGGGGGGCAACACGCTCTGAACCTTATTCCTCAACATGGGCTCCATGAA	858		
Db	793	ATCCAAAGCTTTCGGGGGGCAACACGCTCTGAACCTTATTCCTCAACATGGGCTCCATGAA	852		
Qy	859	AGCATCTTGGAGATCTTATATAAGACAGATATCCAAATTCCTGAGACTGGCTTACCA	918		
Db	853	AGCATCTTGGAGATCTTATATAAGACAGATATCCAAATTCCTGAGACTGGCTTACCA	912		
Qy	919	GAACTGACGCTTGTTCACGAGCCCTAATGTACATTTGATTTGATTCCTGAGCTTAAATGG	978		
Db	913	GAACTGACGCTTGTTCACGAGCCCTAATGTACATTTGATTTGATTTGATTCCTGAGCTTAAATGG	972		
Qy	979	ACTGACCTGATAGAGAGGCTCTCAATAGTTTCCGGTAAAGATATATGTTTCAACGAA	1038		
Db	973	ACTGACCTGATAGAGAGGCTCTCAATAGTTTCCGGTAAAGATATATGTTTCAACGAA	1032		
Qy	1039	GATCGGGTGACAAAGGCCATACAGAAATCAATCTGCCAAGATTAATCTGCGCAAGAA	1098		
Db	1033	GATCGGGTGACAAAGGCCATACAGAAATCAATCTGCCAAGATTAATCTGCGCAAGAA	1092		
Qy	1099	AGACTCAGATCTTTTTCACGCACTGCGACACATCAGACCGCTTAAAGCGAAGAGAG	1158		
Db	1093	AGACTCAGATCTTTTTCACGCACTGCGACACATCAGACCGCTTAAAGCGAAGAGAG	1152		
Qy	1159	ACTTCGGATTAACACAGCAGCAGCTGCGAACAAGAAACAAAGCTGTGCAAAAGAG	1218		
Db	1153	ACTTCGGATTAACACAGCAGCAGCTGCGAACAAGAAACAAAGCTGTGCAAAAGAG	1212		
Qy	1219	AAATAACCTTGGATGCTGATGATACAACTACACTACAGCAAGACGCGTGGCTGTATCAC	1278		
Db	1213	AAATAACCTTGGATGCTGATGATACAACTACACTACAGCAAGACGCGTGGCTGTATCAC	1272		
Qy	1279	TTTCGCTTAGATTTATTTAACTCCCTGTTTAACTCAGACCTTTGGTAAAGTTTGCTCAG	1338		
Db	1273	TTTCGCTTAGATTTATTTAACTCCCTGTTTAACTCAGACCTTTGGTAAAGTTTGCTCAG	1332		
Qy	1339	TTTCAACCTGGGGTAAGTTAGTTGTGTTGAAGAGATGGTGTACCAAGTAAACAAACTT	1398		
Db	1333	TTTCAACCTGGGGTAAGTTAGTTGTGTTGAAGAGATGGTGTACCAAGTAAACAAACTT	1392		
Qy	1399	ATGCGCTTTTTCATCTTCTTGCCTTTGAAGTA	1431		
Db	1393	ATGCGCTTTTTCATCTTCTTGCCTTTGAAGTA	1425		
RESULT 4	ARI52405	1381 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	ARI52405				
DEFINITION	Sequence 5 from patent US 6232527.				
ACCESSION	ARI52405				
VERSION	ARI52405.1				
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1. (bases 1 to 1381)				
AUTHORS	Mehajan, P. B.				
TITLE	Maize Rad2/FEN-1 orthologues and uses thereof				
JOURNAL	Patent: US 6232527-A 5 15-May-2001;				
FEATURES	Location/Qualifiers				
source	1..1381				
	/organism="unknown"				

BASE COUNT	441 a	269 c	346 g	325 t	ORIGIN
Query Match	92.3%	Score 1350.2	DB 6	Length 1381	
Best Local Similarity	99.4%	Pred. No. 0			
Matches 1355	Conservative	0	Mismatches	8	Indels
					Gaps
49	GGCCGCCCCGACCCGGCCACAGCGCCGCGACAGAGATGGGGCTCAAGGGTTTGACGAA	108			
Db	1 CGACCCACGCGCTCCGGCCACAGCGCCGCGACAGAGATGGGGCTCAAGGGTTTGACGAA	60			
Qy	109 CTGCTGGCGGCAATGCGCCCAAGGCATGAAGAGCAAGATTGCGAGCTACTTGGCG	168			
Db	61 CTGCTGGCGGCAATGCGCCCAAGGCATGAAGAGCAAGATTGCGAGCTACTTGGCG	120			
Qy	169 GCGAAATTCGCGCTGACGCGCAGCATGAGATATACAGTTCCTGATTTGAGTGGAAAG	228			
Db	121 GCGAAATTCGCGCTGACGCGCAGCATGAGATATACAGTTCCTGATTTGAGTGGAAAG	180			
Qy	229 ACAGGCATGGAACCTTCACAATGAAGCTGTGAAGTCACTAGTCATTTGCAAGAAATG	288			
Db	181 ACAGGCATGGAACCTTCACAATGAAGCTGTGAAGTCACTAGTCATTTGCAAGAAATG	240			
Qy	289 TTCACGCGGACAAATGAAGTTCTGGAAGCGGGAAATCAAGCAATTTATGTTTGTATGCG	348			
Db	241 TTCACGCGGACAAATGAAGTTCTGGAAGCGGGAAATCAAGCAATTTATGTTTGTATGCG	300			
Qy	349 AAGCCTCTGATATGAGAAACAAGAGCTTGTAAAGATCTCAAAAAGAGATGACGA	408			
Db	301 AAGCCTCTGATATGAGAAACAAGAGCTTGTAAAGATCTCAAAAAGAGATGATGCA	360			
Qy	409 ACCAAGATTCGACTGAGGCGAGTGAAGTGAAGATTAATATCGATTGCAAAAATTGAGC	468			
Db	361 ACCAAGATTCGACTGAGGCGAGTGAAGTGAAGATTAATATCGATTGCAAAAATTGAGC	420			
Qy	469 AAGAGGACTGTAAAGTTCACAAAGGCAACACACGAAGATTGAACGCTATTGAAGCTT	528			
Db	421 AAGAGGACTGTAAAGTTCACAAAGGCAACACACGAAGATTGAACGCTATTGAAGCTT	480			
Qy	529 ATGGGGGTTCCGTGTGTAAGAGCACCTTTCGAAGCAGAAACAAATGTGCAGCCCTTTCG	588			
Db	481 ATGGGGGTTCCGTGTGTAAGAGCACCTTTCGAAGCAGAAACAAATGTGCAGCCCTTTCG	540			
Qy	589 ATAAACGATTAAGGTTCGCTGTGCTTCACAAAGATATGAGACTCCCTACTTTGGGCGT	648			
Db	541 ATAAACGATTAAGGTTCGCTGTGCTTCACAAAGATATGAGACTCCCTACTTTGGGCGT	600			
Qy	649 CCAGGTCCTTCGTCATATTATGATCCCAAGTTCACAAAGAAATFACCTGTGATGCAATT	708			
Db	601 CCAGGTCCTTCGTCATATTATGATCCCAAGTTCACAAAGAAATFACCTGTGATGCAATT	660			
Qy	709 GATGTGGCAAGGTTTTGAGAGAGCTTGAACCTACCATGAGCAGTTCATTTGATTTGTC	768			
Db	661 GATGTGGCAAGGTTTTGAGAGAGCTTGAACCTACCATGAGCAGTTCATTTGATTTGTC	720			
Qy	769 ATTCGTGTGATGTGACTATTGATGATGATCAAAAGTTCGGGGGGCAAAACACTCTCG	828			
Db	721 ATTCGTGTGATGTGACTATTGATGATGATCAAAAGTTCGGGGGGCAAAACACTCTCG	780			
Qy	829 AAACCTATTCTCAACATGGGTCTCATAGAAAGCATCTTGGAGATTTTAATAAGACAGA	888			
Db	781 AAACCTATTCTCAACATGGGTCTCATAGAAAGCATCTTGGAGATTTTAATAAGACAGA	840			
Qy	889 TATCAAAATTCCTGAGGAGCTGGCTTACCAAGAGCTGACAGCTTGTTCAAAGAGGCTAT	948			
Db	841 TATCAAAATTCCTGAGGAGCTGGCTTACCAAGAGCTGACAGCTTGTTCAAAGAGGCTAT	900			
Qy	949 GTGACATTGCAATTTCTGAGCTTAAATGACATGACCTGATGAGAGAGGCTCTCAAGT	1008			
Db	901 GTGACATTGCAATTTCTGAGCTTAAATGACATGACCTGATGAGAGAGGCTCTCAAGT	960			
Qy	1009 TTCCGAGTAAAGATTAATGTTTCAACGAGAGATCGGGGTGACAAAGGCCATAGACAAATC	1068			

Db 961 TTCTGGTAAAGATATGTTTCAACGAGAGTGGGTGACAAAGGCCATAGAGAATC 1020
Qy 1069 AATCTGCCAAGATTAATCGTCCAGAGAGTCCAGTCTTTTCAAGCCACTGCC 1128
Db 1021 AATCTGCCAAGATTAATCGTCCAGAGAGTCCAGTCTTTTCAAGCCACTGCC 1080
Qy 1129 ACCACATGACGACCGCTTAAACGAGAGACCTTGCATTAACAGCAAGCAGCTGCC 1188
Db 1081 ACCACATGACGACCGCTTAAACGAGAGACCTTGCATTAACAGCAAGCAGCTGCC 1140
Qy 1189 AACAGAAACAAAGCGCTGTGGAAGAGAAATATCTTGATGCTGTGATGATCAACTA 1248
Db 1141 AACAGAAACAAAGCGCTGTGGAAGAGAAATATCTTGATGCTGTGATGATCAACTA 1200
Qy 1249 GCACCTACGAAGACGCGGTGGCTGATCATTGCTTATTTACTTCCCTGTTTAA 1308
Db 1201 GCACCTACGAAGACGCGGTGGCTGATCATTGCTTATTTACTTCCCTGTTTAA 1260
Qy 1309 ACTCAGACCTTTGGTAAAGTTTGTCTCATGTTTCAAGCTGGGTAAGTGTGTTTG 1368
Db 1261 ACTCAGACCTTTGGTAAAGTTTGTCTCATGTTTCAAGCTGGGTAAGTGTGTTTG 1320
Qy 1369 AAGACATGCTGTGACCAAGTACAAACTTATGCTGTTTTT 1411
Db 1321 AAGACATGCTGTGACCAAGTACAAACTTATGCTGTTTTT 1363

RESULT 5
AB021666 1354 bp mRNA linear PLN 25-MAR-2000
LOCUS AB021666
DEFINITION Oryza sativa OsFEN-1 mRNA for FEN-1, complete cds.
ACCESSION AB021666
VERSION AB021666.2 GI:4587224
KEYWORDS FEN-1; endonuclease.
SOURCE Oryza sativa cDNA to mRNA.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretioideae; Oryzaceae; Oryza.
1 (sites)
Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and
Sakaguchi,K.
Plant homologue of flap endonuclease-1: molecular cloning,
characterization, and evidence of expression in meristematic
tissues
Plant Mol. Biol. 42 (3), 415-427 (2000)
2 (bases 1 to 1354)
20256470
Kimura,S., Hashimoto,J. and Sakaguchi,K.
Direct Submission
Submitted (21-DEC-1998) Seisuke Kimura, Science University of
Tokyo, Dept. of Applied Biological Science, 2641 Yamazaki, Noda,
Chiba 278-8510, Japan (E-mail:j6498703ed.noda.sut.ac.jp,
Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)
On Apr 17, 1999 this sequence version replaced gi:4062866.
COMMENT
FEATURES
source
1..1354
/organism="Oryza sativa"
/db_xref="taxon:4530"
67..1209
/gene="OsFEN-1"
67..1209
/gene="OsFEN-1"
/note="endonuclease"
/codon_start=1
/product="FEN-1"
/protein_id="BAA36171.1"
/db_xref="GI:4587225"
/translation="MGIKGLTKLADNAPRAMEKQKESYFGRIAVDASMSYOFLL
VGRGMEETLENEGEVYSHLQGFNRTILFAGIKPYVEFGKPPDLKQELAKRY
SKRDATKELEAVEEGDKDAIEKFSKRVKVTKONNESEKRLRLRMGPVVEAPCEA
EACCAALCINDMVAVASSEMDSLTEGAPRLRHMDPSSKRLPVEFEVAVLELE
LTMDPFDLCLTSCDYCDISKIGGDTALKLRGHSIESILENTINKRYPIDPMP

YOEARRLFKEPNVTLIDIPELKWNAPDEGLVEFLVKEGPNODRVTKAIEKIFKANK
SSQGRLESEFRRPVSTSVPLKRDTSSEKPTKAVANKKTGAGGKKK"
polyA_site 1354
/note="17 A nucleotides"
BASE COUNT 415 a 286 c 331 g 322 t
ORIGIN
Query Match 63.4%; Score 927.4; DB 8; Length 1354;
Best Local Similarly 86.3%; Pred. No. 1,2e-218;
Matches 1037; Conservative 0; Mismatches 161; Indels 3; Gaps 1;
Qy 55 CGGCCACCCGCGACAGCGCCGCGAGAGATGGGATCAAGGTTTGAGCAAACTGCTG 114
Db 37 CGCCGCTCGCCGATCCGTGGAGGCGAGATGGGATCAAGGTTTGAGCAAACTGCTG 96
Qy 115 GCGCAATATGCGCCCAAGCGGATGAGAGAGAGTTCGAGAGCTTCTGGCGCCAAA 174
Db 97 GCGCAATATGCGCCCAAGCGGATGAGAGAGAGTTCGAGAGCTTCTGGCGCCGC 156
Qy 175 ATGCCGTCGACGCGAGCATGATATACAGTCTCTGATGATGAGAGAGAGC 234
Db 157 ATGCCGTCGACGCGAGCATGATATACAGTCTCTGATGATGAGAGAGAGC 216
Qy 235 ATGGAACCTTCACAAATGAAGTGTGAAGTCACTAGTATTTGCAAGGAATGTTCAAC 294
Db 217 ATGGAACCTTCACAAAGAGAGCGCGGTGAAGTCACTAGTATTTGCAAGGAATGTTCAAC 276
Qy 295 CGGCAATATGATTAAGTACGAGCGGCAATCAAGCAGTATGTTTGTGTCAGAGCCT 354
Db 277 CGGCAATATGATTAAGTACGAGCGGCAATCAAGCAGTATGTTTGTGTCAGAGCCT 336
Qy 355 CCGATATGAGAAACAAAGAGCTTGTAAAGATATCAAAAGAGATGCAACAAA 414
Db 337 CCGATATGAGAAACAAAGAGCTTGTAAAGATATCAAAAGAGATGCAACAAA 396
Qy 415 GATTCGACTGAGCGAGTGAAGTGAAGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 474
Db 397 GATTCGACTGAGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 456
Qy 475 ACTGTAAGGTCACAAAGAGCAACAGCAAGATGTAAGCGGTATTAAGACTTATGAGG 534
Db 457 ACTGTAAGGTCACAAAGAGCAACAGCAAGATGTAAGCGGTATTAAGACTTATGAGG 516
Qy 535 GTTCTGTTGTAAGAGCACCTTTCGAAGCAGACCAATATGTCAGCTTTCATTAAC 594
Db 517 GTTCTGTTGTAAGAGCACCTTTCGAAGCAGACCAATATGTCAGCTTTCATTAAC 576
Qy 595 GATAAGTGTTCGCTGTTGCTTCAAGATGAGTCCCTTACTTTGGGCTCCACG 654
Db 577 GATAAGTGTTCGCTGTTGCTTCAAGATGAGTCCCTTACTTTGGGCTCCACG 636
Qy 655 TTCTTCTGCTATTTATGATCCAAAGTTCAGAAATACCTGTGATGATGATGATG 714
Db 637 TTCTTCTGCTATTTATGATCCAAAGTTCAGAAATACCTGTGATGATGATGATG 696
Qy 715 GCCAAGTTTGGAGAGCTTGAACCTACACATGAGCAGTTCATTTGTTGTCATCCG 774
Db 697 GCCAAGTTTGGAGAGCTTGAACCTACACATGAGCAGTTCATTTGTTGTCATCCG 756
Qy 775 TGTGATGTGACTATTTGATAGCATCAAGATATGCGGGGGCAACAGCTGGAACCTT 834
Db 757 TGTGATGTGACTATTTGATAGCATCAAGATATGCGGGGGCAACAGCTGGAACCTT 816
Qy 835 ATTGCTCAACATGGGTTCATGAGAAAGCATCTTGGAGATCTTAATAAGCAGATATCA 894
Db 817 ATTGCTCAACATGGGTTCATGAGAAAGCATCTTGGAGATCTTAATAAGCAGATATCA 876
Qy 895 ATTCTGAGGAGCTGCTTACCAAGAGCTGAGCTTGTTCAGAGAGCCTAATGTGACA 954
Db 877 ATTCTGAGGAGCTGCTTACCAAGAGCTGAGCTTGTTCAGAGAGCCTAATGTGACA 936
Qy 955 TTGGAATATTCGAGCTAAATATGACGACCTGATGAGGGGTTCATTAAGTTCCG 1014

Db 937 TTGATATTCCTGAGCTGAAATGCAATGCTCCGATGAGAGAGGTCTTGAGATTTCG 996
Qy 1015 GTAAAGATTAATGCTTTCACAGAGATCGGATGACAAAGAGCCATAGAGATCAATCT 1074
Db 997 GTGAAGAAATGCTTTCACACAGATAGAGTGACAAAGCCATAGAGAAATTAATTT 1056
Qy 1075 GCCAAGATTAATGCTGCGAAGAGACAGCTGAGTCCCTTTTCAAGCCAACTGCCACCA 1134
Db 1057 GCCAAGAAATTAATGCTTTCACAGAGAGAGGTCTTCTTCAAGCCAGTGTAGACCA 1116
Qy 1135 TCAGACCGCTTAAGAGAGAGAGATTCGATTAAGAAACAGAGAGAGAGAGAGAGAG 1194
Db 1117 TCAGAGCTCTGAAAG 1176
Qy 1195 AAAACAAG--GCTGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1251
Db 1177 AAAACAAG 1236
Qy 1252 C 1252
Db 1237 C 1237

RESULT 6
XLU68141 1469 bp mRNA linear VRT 26-JUL-1999
LOCUS Xenopus laevis xFEN1b mRNA, complete cds.
DEFINITION U68141
ACCESSION U68141.1 GI:1549392
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis
Xenopus laevis
African clawed frog.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 1469)
Bibikova,M., Wu,B., Chi,E., Kim,K.H., Trautman,J.K. and Carroll,D.
Characterization of FEN-1 from Xenopus laevis. cDNA cloning and
role in DNA metabolism
J. Biol. Chem. 273 (51), 34222-34229 (1998)
99069415
PUBMED 9852084
2 (bases 1 to 1469)
Bibikova,M., Chi,E., Wu,B., Kim,K.-H. and Carroll,D.
Direct Submission
Submitted (27-AUG-1996) Biochemistry, Univ. of Utah, 50 N. Medical
Drive, Salt Lake City, UT 84132, USA
JOURNAL
TITLE
AUTHORS
REFERENCE
CDS

FEATURES
source
1..1469
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="Hemmati-Brianlou et al. (1991) Development
111, 715-724"
/dev_stage="tadpole"
/note="also present in oocytes, as determined by
sequencing of RT-PCR products"
120..1268
/function="5'-3' exonuclease and a structure-specific
endonuclease that removes 5' single-stranded 'flaps';
implicated in processing of Okazaki fragments during DNA
replication and in base excision repair; 5' nuclease"
/note="FEN-1 homolog; similar to mammalian Dnaase IV or
MFI, to S. cerevisiae RAD27, RTH1 or YKL510, and to S.
pombe rad2; similar to Xenopus laevis xFEN1a encoded by
GenBank Accession Number U64563"
/codon_start=1
/product="xFEN1b"
/protein_id="AAB08478.1"
/db_xref="GI:1549393"
/translation="MGIRGLAKLADVAIPAIEKHDISKYFGKAVVDASMCIOPLI
AVRODGNMLONEGETTSHLMGMFRTIMLEHGKIPYVFGPKPOKSEELAKRSE
BRAAEKLEAAEEAGEVENIEKFKRLVYTKVKNNECKLLSLMGIPIYDACEAF

ATCAALVKAQVYAATEDMDALPEPTVLLRHLLTASEAKRLPIQEFHLNRVODIGI
NHQEVDFDILGSDYCEITRIGIRKRAIDIRHKTLEIIDNIDIKYPIDENMLH
KEAROLFLEPEVIDADITELKTEDEGLVAFPMKGEQFSEDI RNQAKKLAKNRQG
STQGRIDDFEKFYTGSSISKRREVSRSKTSKSTGTGPGRKRKX"
BASE COUNT 461 a 288 c 377 g 343 t
ORIGIN
Query Match 22.3%; Score 326.8; DB 5; Length 1469;
Best Local Similarity 58.1%; Pred. No. 5e-70;
Matches 615; Conservative 0; Mismatches 437; Indels 6; Gaps 2;
78 AGAGCAATGGGATCAAGGTTTGACGAAGTGCCTGGGAGCAATGCGCCCAAGCCAT 137
Db 113 AGTCAATATGGGAATTCACGGCTTGCCCAAGCTTATCGAGATGTGGCCCTGCAGCTAT 172
Qy 138 GAAGAGCAAGATTCGAGAGACTTCCGGCGCAAAATCGCCGTGAGCGCCAGCATAG 197
Db 173 CAAGAGCAATGATATAAAGATTCTTTGGTGAAGAAAGTGGCGGTGATGCTTATGTG 232
Qy 198 CATATACCATGCTCTGATTGTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 257
Db 233 CATTTACCATGCTCTTATGCTGCTGACAGAGATGGA--ATATGCTGACAGATGAGA 289
Qy 258 TGGTGAAGTCAATGATTCGACAGAAATGTTCAACGAGCAATTAAGATTACTGGAAGC 317
Db 290 AGGCGAAGACCAAGAGCATCTCATGGGTATGTTTACGCTACCATTTGCGATTGGAGCA 349
Qy 318 GCGAATCAAGCCGATTATGTTTATGATGCAAGCCCTCTGATATGAGAAACAAGACT 377
Db 350 TGGCATCAAAACCGATTATGTTTATGTTGATGGAAGAGAGAGAGAGAGAGAGAG 409
Qy 378 TGCATAAGATGATCTCAAAAGAGATGATGATCAACCAAGATGATGAGAGAGAGAG 437
Db 410 AGCCAAAGCAGATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 469
Qy 438 AGGAGATTAAGATGCGATTGAAAAATTTGAGCAAGAGAGACTTAAAGGTGACAGGCA 497
Db 470 TGGGAGAGTGAAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Qy 498 CAACGAGATTTGAACGGCTATTAGAGCTTATGGGGGTCTCTGTTAGAGAGAGAGAG 557
Db 530 CAATGAAGATGAG 589
Qy 558 TGAAGCAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 617
Db 590 TGAAGCAG 649
Qy 618 AGAGATATGAG 677
Db 650 TGAAGATGAG 709
Qy 678 AAGTCCAG 737
Db 710 TGAAGATTAAG 769
Qy 738 ACATCAGATGAG 797
Db 770 CATCAATCAG 829
Qy 798 CATCAAGATGATGGGGGCAACAGCTCTGAAGACTTATGCTCAACATGAGGATCAGAG 857
Db 830 AATCCGATGAG 889
Qy 858 AAGCATTTGAG 917
Db 890 AGAGATCATGATTAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 949
Qy 918 AGAGATCGAG 974
Db 950 GAGAGGCGAG 1009
Qy 975 ATGAGCTGAG 1034

Db	Sequence	Length
Db 1010	GTGGACTGAGACCCTGGATGAGAGAGGCGACTGCTGGCCCTCATGTGTTGGTGGAGCAAGCAAGTTCAC	1069
Qy 1035	CGAAGATCGGGGTGACCAAAAGGCCCATGAGCAAGTCAAAATGTGGCCCAAGATATTAATCTGTGCA	1094
Db 1070	TGAAGACCGGATTAACGCAATATGGGGCCAAAABTTGGCCAAAGACCGTCACAGCGACGACGCA	1129
Qy 1095	AGGAACTCGAGTCTCTTTTCAACGCCACATGCCACCA	1132
Db 1130	GGGCGCGCTGGATGACTTTTTCAAAGTAGCTGAGATCCA	1167

RESULT 7				
LOCUS	XL064563	1461 bp	mrna	
DEFINITION	Xenopus laevis 5' nuclelease xRNA		linear	
ACCESSION	nc0562		VRT 26-JUL-1999	
				complete cds.

SOURCE	ORGANISM
African clawed frog.	Xenopus laevis

Ampipruræ, Baliracnidae; Anura; Mesobatrachidae; Pipridæ; Pipridæ; Xenopodinae; Xenopus.

J. Biol. Chem. 273 (51), 34222-34229 (1998)
99069415
JOURNAL
MEDLINE

FEATURES
 ABSTRACTING
 AUTHOR INDEX
 TITLE
 JOURNAL
 Submitted (19-JUL-1996) Biochemistry, U. of Utah Med. Sch., 50 N
 Medical Dr., Salt Lake City, UT 84132, USA
 Location/Qualifiers

```

/organism="Xenopus laevis"
/db_xref="taxon:8355"
111.1259
CDS

```

```

implemented as processing unit.
during DNA replication and in base excision repair;
specifically removes single-stranded 5' sequence
/note="similar to FEN-1, DnaseIv, MFI, RAD27, RTH1, and
rad2"
/codon_start=1
/product="5, nuclease xFenia"
/protein_id="AAB06176.1"
/db_xref="GI:1490870"
/translation="MGHIGLKLIADYAPAIAIKENDIKSYFGKRYAVASMCIOYFLL
AVROGNTLONDEGETTSHLMGMYRTIRAVEGIGIPVYFQDKPOMKSGCLARSGLAR
RAAEKELLEAAEEAGEVENTEKTRKLYVYKONHDEKCLITLNGIIRVYAPAEADIG
ATCALVAKGAVYAAATFEDMDALTEGPVLLRHLTA SPEAKKLPIQEFHINRYODIG
THEQYVDCILILGSDYCEITIGISPKRAIDLIQHTIDEILINDIDLUKYPYPEWMLL
SKAGHLLEPEVVDITELKWLIEPEDEGSVAIMCEKQSEDIRNGKAKLIANKRQY
KQKRLIDFFEVATGVSSTKRKEESGSAKKRAKAGTGPACFKRKG"
BASE COUNT      454 a      293 c      365 g      349 t
ORIGIN

```

	Query Match	21.1%	Score 308.6	DB 5	Length 1461
	Best Local Similarity	55.7%	Pred. No. 1.6e-65		
	Matches 633	Conservative	0	Mismatches 499	Indels 6
				Gaps 2	
QY	85	ATGGGCATCAAGGCTTTGACGAACATGCTGGCGGACAATGCGCCCAAGGCATGAMGAG	144		
DB	111	ATGGGAATTCACGGTTTGGCCAAACTTATTCGACAGATGTGGGCACTCGACGATCAACAAAG	170		
OY	145	CAGAAAGTTCAGAGACTACTTCGGCGCAAAATTCGCGCGTCGACGCGACATGAGCATTTAC	204		
DB	171	CATGATTTAAAAAGTTACTTTTGTGTGCAAAAGTGGCGGCGTCAATGCTTCCATGTGCATTTAT	230		
OY	205	CAGTTCCTGATTTAGTTAGTGGAMGGACGCAATGAACTCTTCACAAATGAGAGCTGGTGA	264		

[illegible]

/db_xref="taxon:8355"
/cell_type="ocyte"
1.1149
/gene="FEN1"
1.1149
/gene="FEN1"
/function="structure-specific nuclease"
/product="flap endonuclease 1"
/protein_id="AAB8707.1"
/db_xref="GI:2674207"
/translational="MGIHLAKLADYAPAIKHEHDKSRERKVAIDASMCIOFLI
AVRODNTIONEGCTTSHLMFRTITMVEGKIPYVEDEKRPQKSKSLARSE
RRBAEKLLEAEAGSEVENIEKFTRLVKTROHNEBECKRLTTLGLPYDAPEAE
ATCAALVKAKVAAATIEDMALFETGPLRLHLAASEKKLPIDFHLNRYIDIGI
THRDYDLCILGSDYCEITRIGIKRPAIDLIROKHTIDEIIDNIDLKVPYENMLH
KEAKHLEPEVVDITTELKWIPEDECEIVAFMCEKOFSDRIENGKRIAKNRQG
STOGRLLDFPKVTVGSVSTRKRESEKSKAKKAKTGTGPAPKFRGK"

BASE COUNT 352 a 247 c 306 g 244 t
ORIGIN

Query Match 20.9%; Score 305.4; DB 5; Length 1149;
Best Local Similarity 55.5%; Pred. No. 9,4e-65;
Matches 632; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 85 ATGGCATCAGAGGTTGACGAACCTGCTGGGCGACATCGGCCCAAGGATGAAGAG 144
DB 1 ATGGCAATTCACGCTTTGGCCAAACTTATCGCAGATGTGCACCTGCAGTATCAAGAG 60
QY 145 CAGAGTTGAGAGCTACTTGGCCGCAAAATCGCTCGACAGCCCATGACCATATAC 204
DB 61 CATGATATATAAAGTTACTTGTGCGAAAGGTGCGATCGATCGCTTCATGCTATAT 120
QY 205 CAGTTCCTGATTTGATGTTGGAAGACAGCATGGAACCTGCACAAATGAGCTGTGAA 264
DB 121 CATGTTCTTATTTGGGTGAGACAGAGTGGCA--ACACGCTGCAAAAGCAAGAGTAA 177
QY 265 GTCACTAGTCTTTCAGCAAGATGTTCAACCGGACAAATGATTAAGTACTGGAAGCGGAGAT 324
DB 178 ACCACAGCCATCTTATGGTATGTTTACCGTACCATTCGCATGATGATGAGCATMGATC 237
QY 325 AAGCAGTTTATGTTTATGATGAGCAAGCTCTCTATATGAAGAAACAAGCTTCTTAA 364
DB 238 AAACCAAGTATGTTTATGATGAGCAAGCTCTCTATATGAAGTCCGTTGAACTTAA 297
QY 385 AGATACCTCAAAAAGAGATGATGACCAACCAAGATGATGATGAGAGTGAAGAT 444
DB 298 GCGAGTGAAGCGTGAAGGCGGAGAGAGAGTGTGAGTGGCGGAGAGGCTGGAGAA 357
QY 445 AAAGATCGGATTTGAAAATTTAGCAAGAGCACTGTAAAGTTCACAGGACACACAGAA 504
DB 358 GTGGAGAACATTGAGAACTTCAAAAAGCTTCTCAAGGTACCAAGAGCAACAATGAG 417
QY 505 GATTTAAACGGCTATTAAGACTTATGGGGGTCTCTGTTAGAGGACCTTCTGAAGCA 564
DB 418 GAGTCAAGAAATTAATTAACCTTAATGGGATTCCTTATGATGATGATGATGATGAT 477
QY 565 GAAGCAAGATGTCAGCCCTTTCGATTAACATTAAGTGTTCGCTGCTTCAAGAT 624
DB 478 GAGGCGACGTGTGCTTGTAGTAAAGCAGGAAAGTGTGATGATGATGATGATGAT 537
QY 625 ATGGATCCCTTACTTGTGGGGCTTCACAGGTTCTGCTCATTAATGATGATGATGAT 684
DB 538 ATGGATGCCCTTGTGACCTTGTGACACCCCTGTGCTTATGAGCACTTCACAGCAAGGAG 597
QY 685 AAGAAATACCTGTGATGGAATTTGATGTTGCAAGGTTTGGAGAGCTTGAAGTACC 744
DB 598 AAAAGGCGCTTATTCAGAGAGTTCATCAACCGTTCATTCGACGACATAGGATACCA 657
QY 745 ATGAGCAATTCATGATTTGATGATCCTGTGATGATGATGATGATGATGATGATGAT 804
DB 658 CATGAAACATTTGTGATCTCTGATATCTTGTGAGAGTGAAGTGAAGTGAAGTGAAG 717

QY 805 GGTATCGGGGCGCAACAGCTCTGAACCTTATTCGCAACATGGTGCATAGAACATC 864
DB 718 GGCATGTGACCCCAAAAGAGCGCATTTGACCTGATCGCTAGCATTAAGACATTTGACGATC 777
QY 865 TTGAGAAATCTTAAATAAGACAGATATCAATTCCTGAGAGCTGCTTACCAAGAGCT 924
DB 778 ATGATTAACATTTGACCTCAAGAAATTAACCCCTGCGGAAACCTGGTGCACAGAGGCA 837
QY 925 CGAGCTTGTTCAGAGAGCTTAA--TGTCACATTTGATATTCCTGAGTAAATAGACT 981
DB 838 AAACATCTTTCTTTCAGAACGAGAGTGTGACATGACATGACATGACGAGCTGAAGTGAATC 897
QY 982 GCACCTGATGAGAGGCTTCATAGTTTCTGTTAAAGATTAATGTTTCAACAGAT 1041
DB 898 GAGCCCGATGAAGAGGATTTGTGCTTCATGTTGTTGAGAGAGAGATTAATGATGAAGT 957
QY 1042 CGGGTGAACAGCCCATAGAGAGATCAAAATTCGCCAAGATTAATTCGCAAGAGAA 1101
DB 958 CGCATACCCCAAGGCGCCAAAGTTGGCCAGAGACCGGCAAGCAGCAGCAGGCGCG 1017
QY 1102 CTGAGTCTCTTTTCAAGCCACTGCCACCATCATGACACGCTTAAACGGAAGAGACT 1161
DB 1018 CTGATGACTTTTCAAGTGCACCGGATCCGTTAGCTCAACTTAAAGAAAGAGCAGAA 1077
QY 1162 TCGATTAACACAGCAGGCGAGCTGCGACACAGAAACAAAGGCTGTGGAAGAGAA 1220
DB 1078 TCCAAAGATCCGCTTAAAGAAAGCAAAACCTGAGAGTACCCCTGCAGGAAAGTTCAA 1136

RESULT 10
LOCUS BC010203 2163 bp mRNA linear ROD 22-OCT-2001
DEFINITION Mus musculus. Similar to flag structure-specific endonuclease,
ACCESSION BC010203 IMAGE:3158176, mRNA, complete cds.
VERSION BC010203.1 GI:16307327
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jelifire Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center
Center code: BCK-HSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villaloboscm.tmc.edu,
Villalobos, D.K., Luna, R.A., Hale, S.M., Huilyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny,D.M., Gibbs,R.A.

FEATURES
source
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAP Plate: 6 Row: d Column: 19.
Location/Qualifiers
1. 2163

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:7137 IMAGE:3158176"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."

```

/clone_1lb="NCI_CGAP_Mam6"
/lab_host="DH108"
/notes="Vector: PCMV-SPORT6"
206..1348
/codon_start=1
/product="Similar to flag structure-specific endonuclease"
/protein_id="AAH10203.1"
/db_xref="GI:16307328"
/translation="MGHGLAKLADYAPSAIRENDIKSYFRKRAIDAISMSIYFLI
AVRGGVLDONAEEMEEVEKFKRLVYVTFKOHNECKHLSIPDLAPSAE
RRARAEKOLDAEEMEEVEKFKRLVYVTFKOHNECKHLSIPDLAPSAE
ASCAALAKKAVYAAATBEDMCLTFGSPVIMRLTASEAKLLEIOEPHLSRVLOELG
NOEOPVLCILSDICESIRIGPRAVYIQRKHSIEITVRLDPSKPYVENMLH
KEAOFLPEPEVDPESEVELKMSPEEELVTKGKQSEERISVYKRLSKSHOG
STOGRLDPEFVVGSLSSAKRKEPEPKRPAKKRAKTGGAKFRFRGK"
BASE COUNT      540 a      528 c      619 g      476 t
ORIGIN
Query Match      20.9% Score 305.4; DB: 10; Length 2163;
Best Local Similarity 56.3%; Freq. No. 1.1e-64;
Matches 642; Conservative 0; Mismatches 481; Indels 18; Gaps 3;
QY      85  ATGGGATCAAGGTTTGGACGAAGTCTGGCGGACAAATGGCGCCAGCGATGAAGAG 144
      ||||| || || || || || || || || || || || || || || || || || || ||
Db      206  ATGGGAATTCAGGGCTTGCCAACTAATGTGTGATGTGGCCCCAGGCGCATCGTGAG 265
QY      145  CAGAACTTGAGAGCTACTTGGCGCAAAATGCCGCTGGACGCCATGACATATAC 204
      || || || || || || || || || || || || || || || || || || || || ||
Db      266  AATGACATCAAGAGCTACTTGTGGCAAAAGTGGCCATGATGCCATGACATCTAC 325
QY      205  CAGTCCCTGATGTGTGGGAAGGACGATGAAACTCTGCAAAATGAAGCTGTGTAA 264
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      326  CAGTTCCTATTTCTGTGTGTGAG--GTTGGGATGTGCTCAACAGAGAGAGGTGAG 382
QY      265  GTCACTAGTCATTTGCAAGAAATGTTCAACGACAAATAGATTAAGTACTGGAGCGGAATC 324
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      383  ACCACGAGCACTGATGTGGCATGTTCTACACCATCCGATGATGAGAAATGGCATC 442
QY      325  AAGCCAGTTTATGTTTGTATGCAAGCTCTGATATGAAGAACAGAGTGTCTAA 384
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      443  AACCTGTGACTCTTTATGCAAAACACACACACTGAAGTCAAGGCGGCTGGCCAG 502
QY      385  AGATACTCAAAAAGAGATGATGCAACCAAGATCTGACTAGGACGATGAGTAGAGAT 444
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      503  CGCAGTGAAGAGGCGCGCCGACGAGTGAAGAACATCAGACAGGCTGAGAGGCTGGATG 562
QY      445  AAGATGCGATTGAAAAATGAGCAAGAGACTGTAAAGTCAAGCAAGCAACAGAA 504
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      563  GAGGAGGAGGTGAGAAAGTTCAACAGAGGCTGTGAAGGTCAACCAAGCAACATGAT 622
QY      505  GATTGTAACGGCTATTAGACTTATGAGGGGTCTGTTAGAGGACCTTCTGAGCA 564
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      623  GAGTGAACACCTGCTGAGACCTCATGAGGATCTCTTACTTGATCACCAGCAGAGGA 682
QY      565  GAAGCAAAATGTGACAGCCCTTTGCATAAAGATAGGTGCTGCTGCTTCAGAAAT 624
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      683  GAGGCCAGCTGTGCTGGCCGCAAAAGGCTGCAAAAGTATCTGCGGCACGAGAGAC 742
QY      625  ATGACATCCCTTACTTTTGGGGCTCCAGGTTCTTCGTATTAATGATGATCAATGCC 684
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      743  ATGACATGCTCTCACTTTGACAGCCCGTCTATGCTGACACTTAAGTCACTGAGTGGCC 802
QY      685  AAGAAATACCTGATGATTTGATTTGATTTGCAAGTTTGGAGAGCTTGAACCTCAC 744
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      803  AAGAACTGCTCATCAAGAGTTCCATCTAGCGCGCTCTGACGAGAGTGGGTCTTGAA 862
QY      745  ATGACACCAATTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 804
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      863  CAGAGAGAGTTTGTGATCTGTGATCTGTGATCTGTGATCTGTGATCTGTGATCTGT 922
QY      805  GGTATCGGGGGGCAAGAGCTGAACTTATTCGTAACATGGGTCCATGAAGAACATC 864
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      923  GGCATTTGGGCCCAAGCGGCTGTGATCTCAAGAAACATTAAGACATGAGAGATC 982

```

```

QY      865  TTGGAAATCTTAATTAAGACAGATATCAATTCGAGGAGCTGGCTTACCAGAACT 924
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      983  GTGAGCGGCTGGACCCGACAGTACCCGTTCCAGAACTGCTCCACAGAGAACCC 1042
QY      925  CGACGCTGTTCAAGGAGGCTTAATG---TCACATTGATATTCCTAGACTAAATGACT 981
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1043  CAGCAGCTCTCTCTGAGCAGAGTACGAGACCCAGATCTGTGAGACTGATGAGAGC 1102
QY      982  GCACCTGATGAGAGGCTCTCATAGTTTCCTGTTAAAGATTAATGTTTCAAGCAAGAT 1041
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1103  GAGCCAAATGAAGAAAGTGTGTCATTAATTTATGTGTGTGAAGAACAGTTTCTGAAG 1162
QY      1042  CGGGTCAAAAGCCATAGCAAGATCAATATCTGCCAAGATTAATTCGCAAGAGAGA 1101
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1163  CGAATTCGACAGTGGGTCAAGGCGCTGATGAGAGCGCGCAGGAGCAGCAGGAGAGCC 1222
QY      1102  CTCGAGTCTTTTTCAGCCCACTGCCACATCAGCACCCCTAAACGGAAGAGACT 1161
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1223  CTCGATGATTTTTCAGAGGAGCAGGCTCCTCTCTCAGC-----TAAGCGC 1270
QY      1162  TCGGATTAACCAAGCAAGCAGCTGGACCAAGAAAGAGGCTGTGGAAGAGANA 1221
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1271  AAGAGCCAGAACCCAGAGGCGCTGCTAAGAGAAAGCAAAAGACTGGGGAGCGGGAAG 1330
QY      1222  T 1222
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1331  T 1331

RESULT 11
AC026761      166088 bp      DNA      linear      HTG 25-JAN-2002
LOCUS      Mus musculus chromosome 19 clone RP22-325J22 strain 129/SVETAC1br,
DEFINITION      WORKING DRAFT SEQUENCE, 25 unordered pieces.
ACCESSION      AC026761.10 GI:15217172
VERSION      AC026761.10 GI:15217172
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      house mouse.
ORGANISM      Mus musculus.
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 166088)
Grills,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Fereta,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
Direct Submission
Submitted (24-MAR-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
On Aug 21, 2001 this sequence version replaced gi:10937955.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site:
http://wehanning.bwh.harvard.edu:9088/hpccg/jsp/hpccg/sequence/mous
e.html
Contact: gnktm@capecod.bwh.harvard.edu
-----Summary Statistics
Center project name: ACD
Sequencing vector: pUC18; 108752
Chemistry: Dye-terminator Big Dye 100%
*Consensus quality: 156712 at least 920
*Consensus quality: 152694 at least 930
*Consensus quality: 147247 at least 940
Estimated insert size: agarose-FP - N/A
**Estimated insert size: 165608 - sum-of-contrigs

```



```

Db 11711 CGCAGTGAAGAGCGCGCCGAGAGCTGAGAAAGCACTGCAGCAGGCTCAGAGCTGGGATG 11770
QY 445 AAGATGCGATTGAAAAATTTGAGCAGAGAGCACTGAAAGTGCACAAAGGACACAGCAAGAA 504
Db 11771 GAGGAGGAGGTGGAGAAATTCACAGAGAGGCTCGTGAAGGTGCACCAAGCAACACATATGAT 11830
QY 505 GATTGTAAACGGCTATTAAGACTTATGGGGGTCTCTGTTAGAGGCACTCTTCTGAAGCA 564
Db 11831 GAGTCAAAACACTCTGCTGAGCCTCATGGGCAATCCCTTACTTATGACACCCAGGAGAGCA 11890
QY 565 GAAGCAGATGTGCAGCCCTTTGCAATAAGCATAGAGTGTTCGCTGTCTTGAAGAT 624
Db 11891 GAGGCCAGCTGTGCTCCCTGGCAAAAGCTGGCAAGTCTATGCTGCGGCAAGAGAGAC 11950
QY 625 ATGAGCTCCCTACTTCTTTGGGGGCTCCAGGTTCTCTGCTCAATTAATGATCCAAATTC 684
Db 11951 ATGAGCTGCTCACTTTTGGCAGGCCCCGCTAATGCGCACTTAAGTCCAGTGAGGCC 12010
QY 685 AAGAAATACCTGTGATGCAATTTGATGTCGCAAGGTTTGGAGGACTTGAATCACC 744
Db 12011 AAGAAAGCTCCCATCCAGAGAGTTCATCGAGCCGCTCCAGAGAGCTGGGTCTGAAC 12070
QY 745 ATGAGCACTCATTTGATTTGTCATCCCTGCTGATGATGATGATGATGATGATGATGATGAT 804
Db 12071 CAGGAGCACTTTGTGATCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12130
QY 805 GGTATCGGGGGGCAAAACAGCTCTGAAACTTATTCGTCACATGAGTGGTCCATAGAAAGATC 864
Db 12131 GGTATTTGGGCCCAAGCGGGCTGTGGATCTCATCCAGAAACATTAAGAGCATCGAGAGATC 12190
QY 865 TTGGAGATCTTAATTAAGACAGATATCAAAATTCCTGAGAGACTGGCCTTACCAAGACT 924
Db 12191 GTGAGCGGGCTGAGACCCCAAGAGTACCCGTTCCAGAGAGCTGCTCCAGAGAGGCC 12250
QY 925 CGAGCTGTTGTCAGAGAGCTTAATG---TCACATTGATATTCCTGAGAGTAATGAGACT 981
Db 12251 CAGCAGCTCTCTCTGAGAGCAGAGATGAGAGCAGACAGCTCTGTGAGCTGAAGTGGAGC 12310
QY 982 GCACCTGATGAGAGAGGTCTCATAGTTCTCTGTTAAAGATATGTTTCAACGAAAT 1041
Db 12311 GACCAAAATGAGAGAGTGTGTCAAATTTATGTTGTAAGAAAGCAGTTTCTGAAGAG 12370
QY 1042 CGGGTGCACAAAGGCCATAGAGAGATCAATCTGCCAAGATTAATCTGCCAAGAGAGA 1101
Db 12371 CGAATTCGAGTGGGTCAAGCGCTGAGTGAAGCGCCGAGGGCAGCAGCAGAGAGGCC 12430
QY 1102 CTCGAGTCTTTTTCAGAGCACTGCCACACATCAGACCGCTAAACGGAAGGAGACT 1161
Db 12431 CTCGATGATTTCTTCAAGGTGACAGGCTCAGCTCTCTCAGC-----TAAAGCGC 12478
QY 1162 TCGGATTAACCAAGCAAGAGAGCTGCGAACAAGAAACAAAGGCTGGTGAAGAAAGAA 1221
Db 12479 AAGGAGCCAGAACCCCAAGGGGCTGCTGAAGAAAGCAAGACTGGGGAAGCGGGAGAG 12538
QY 1222 T 1222
Db 12539 T 12539

```

```

RESULT 12
HSRAD2 1757 bp mRNA linear PRI 07-DEC-1998
LOCUS Homo sapiens mRNA for flap endonuclease-1.
DEFINITION X76771.1 GI:1905802
VERSION X76771.1
KEYWORDS Ien1 gene; flap endonuclease-1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1757)
AUTHORS Watts,F.

```

```

TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) F. Watts, University of Sussex, School of
Biological Sciences, Falmer, Brighton BN1 9QG, UK
REFERENCE 2 (bases 1 to 1757)
AUTHORS Murray,J.M., Tavassoli,M., al-Harthly,R., Sheldrick,K.S.,
Lehmann,A.R., Carr,A.M. and Watts,F.2.
TITLE Structural and functional conservation of the human homolog of the
Schizosaccharomyces pombe rad2 gene, which is required for
chromosome segregation and recovery from DNA damage
JOURNAL Mol. Cell. Biol. 14 (7), 4878-4888 (1994)
MEDLINE 94277093
FEATURES
source Location/Qualifiers
1..1757
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 368..1510
/gene="Ien1"
CDS 368..1510
/gene="Ien1"
/codon_start=1
/product="flap endonuclease-1"
/protein_id="CAA54166.1"
/db_xref="GI:3980293"
/translation="MGIGLAKLIDVAPSAIRENDIKSYEGRKVAIDASISYFELI
AVROGDDVLQNEGETTSHLMGMFYRTIRMMNGIRPYVDPGKPOLSKDELRSE
RRAAEKROLQAOAAGAEQVEKFRRLVKVTRKHNDCKHLISLMGIPYDAPSEAR
ASCAALVRKGVYAAATEDMDCLTFGSPVLMHLTASAEKKLPIQEFHSLRLOELG
NOROPVDILIGSDVCEISIRIGTGPRAVDLQKHRSIEIYRDLDPNKPYPENMLH
KEHOLFLEPEVLDPESVCEIPLKWSPEPEELIKFEMGGEFOFSERISRVKRLSKSROG
STGVRDIDDFKVTGSLSSAKRKEPEKSGTTRKKATGAAGKTRKK"
BASE COUNT 418 a 443 c 542 g 354 t
ORIGIN
Query Match 20.7%; Score 302.8; DB 9; Length 1757;
Best Local Similarity 55.08; Pred. No. 4,5e-64;
Matches 661; Conservative 0; Mismatches 532; Indels 9; Gaps 3;
QY 85 ATGGGCATCAAGGGGTTTGACAAACTCTGCGGCAATCGCCGCCCAAGCGCATGAAGAG 144
Db 368 ATGGGAATTCAGAGCCCTGGCCCAACTAATTCATGATGTGGCCCCCAGTGCATCCGGAG 427
QY 145 CAGAAATTCGAGACTACTTGGCGGCAAAATCGCCGTCAGAGCGCATGAGATATAC 204
Db 428 AATGACATCAAGACCTCTTTGGCGGTAAAGTGCCATGATGCTTATGACATTTAT 487
QY 205 CAGTCTGATTTGATGTAAGGACAGGATGGAATCTCACAAATGAAGCTGTGAA 264
Db 488 CAGTTCCTGATTTGCTGTTGGCCAG---GTGGGATGATGCTGACGATGAAGAGGTGAG 544
QY 265 GTCATAGCTATTTGCAAGATGTTACCGGACATATAGATTACTGGAAGCGGAATC 324
Db 545 ACCACACAGCCACCTGATGGGATGTTTACCGCACCATTCGCATGATGAGGAACGGCATC 604
QY 325 AAGCAGTTAATGTTTGTGATGGCAAGCTCTCTGATTTGAAGAAACAAGAGTGTCTAAA 384
Db 605 AAGCCGTGTATGCTTTGATGAGCAGAGCGCCACAGCTCAAGTCAAGGAGAGCTGGCCAA 664
QY 385 AGATACCTCAAAAAGATGATGACCAAGATCTGACTGAGGAGTGAAGTGGAGAT 444
Db 665 CGCAGTGAAGCGGGGCTGAGGAGAGAGAGCAGTGCAGCAGGCTCAGGCTGAGGGGCC 724
QY 445 AAGATGCGATTGAAAAATTTGAGCAGAGAGCTGAAAGTGCACAAAGGACACACAGAA 504
Db 725 GAGCGAGGAGGTGAGAAATTCATTAAGGCGGTGGTGAAGGCTCAAGCAGCAAAATGAT 784
QY 505 GATTGTAAACGGCTATTAAGACTTATGGGGGTCTCTGTTAGAGGCACTCTTCTGAAGA 564
Db 785 GAGTCAAAACATCTCTGAGCCTCATGGGCAATCCCTTATCTGATGACACCCAGTGAAGCA 844
QY 565 GAAGCAGATGTGCAGCCCTTTGCAATAAGATTAAGTATGATGCTGCTTCTGCAAGAT 624
Db 845 GAGGCCAGCTGTGCTCCCTGTGTAAGCGCTGCAAAAGTCTATGCTGCGGCTACGAGAGAC 904

```



```

OY 625 ATGAGCTCCCTTACTTTTGGGGCTCCACGGTTCTTCTGTCATTTAATGATCCAAATTCC 684
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ATGGACTGCTTCTTCTGAGCCCTTGCTGAATGACACACCTGACCGAGAAAGC 964
OY 685 AAGAAATACCTGTGATGGAATTTGATGTTGGCAAGGTTTGGAGAGCTGAAGCTACC 744
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 AAAAAGCTGCAATCGAGAATTCACCTGAGCCGATTTCTGCAAGCTGGCTTGAC 1024
OY 745 ATGACCACTTCTTATTTGATCTGTGTGATGTGACTATTTGATAGCATGAAA 804
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1025 CAGAAACAGTTGTGTGATCTGTGCAATCTGTAGGCACTGCTGAGAGATGATCCGG 1084
OY 805 GGTATCGGGGGGCAACACCTCTGAACCTTATTCCTCAACATGGCTGAGAAAGCATC 864
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1085 GGTATTTGGGCGCAAGGGCTGTGAGCTCATCCAGAAAGCAAGAGCATGAGAGATC 1144
OY 865 TTGGAGATCTTATTAAGAAGATATCAATTCCTGAGAGCTGGCTTACCAAGAACT 924
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1145 GTGGGGGAGCTGAGCCCAAGATACCTGTGCGCAAAATTTGGCTCCACAGAGAGCT 1204
OY 925 CGAGCTTTGTCAGAGAGCTATG--TCACATTTGATATTTCTGAGCTAAATGAGACT 981
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 CACAGCTCTTTTGAAGCTGAGGTGCTGACCCAGAGTCTGTGAGCTGAGAGTGGAGC 1264
OY 982 GCACCTGATGAGAGAGCTCTCAATCTTCCGTGTAAGAAAGATATGTTTCAAGAGAT 1041
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1265 GAGCAATGAGAGAGAGCTGATCAAGTCTGTGTGTGTAAGAACAGCTTCTGAGAG 1324
OY 1042 CGGTGCAAAAGGCAATAGAGATCAATTCGCCAAGAAATTAATTCGTCGCAAGAGAA 1101
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1325 CGATCCGCGAGTGGGGTCAAGAGGCTGAGTAAGAGCCGCGCAAGGAGCAGGCGCG 1384
OY 1102 CTGAGAGCTTTTCAAGCACTGCGCAC--CACATCAAGCAGCGCTAAAGGAGAGG 1158
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1385 CTGATGATTTTCTTCAAGGTGACCGGCTCACTCTCTTCAAGTCAAGGAGAGCCAGAA 1444
OY 1159 ACTTGAGATTAACAAGCAAGGCTGCGAACAAGAAACAAAGGCTGTGTGAAAGAG 1218
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1445 CCCAAGGATCCACTAAGAAAGAAAGTGGGGGAGGAGAACTTTAAAGAGGGA 1504
OY 1219 AATATCTTGATGCTGTGATGATCAACTAGCACTAGCAAGACAGCGGTGCGATGATC 1278
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1505 AATTAATGTGTTTCCCATTAATACCTCTTCAACCCCAAGAAATTTGGCGTCTTTACCC 1564
OY 1279 TT 1280
      ||
Db 1565 TT 1566
      ||

RESULT 13
BC000323          2031 bp      mRNA      linear      PRI 12-JUL-2001
LOCUS             Homo sapiens, flap structure-specific endonuclease 1, clone
DEFINITION        BC000323
ACCESSION         BC000323
VERSION           BC000323.1 GI:12653112
KEYWORDS          MGC.
SOURCE            human.
ORGANISM          Homo sapiens
                  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 2031)
AUTHORS           Strausberg, R.
TITLE             Direct Submision
JOURNAL           Submitted (15-NOV-2000) National Institutes of Health, Mammalian
                  Gene Collection (MGC), Cancer Genomics Office, National Cancer
                  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                  USA
REMARK            NIH-MGC Project URL: http://mgc.nci.nih.gov
                  Contact: MGC help desk
                  Email: cga@nci.nih.gov
                  Tissue Procurement: DCTD/DTF
                  cDNA Library Preparation: Rubin Laboratory

```

```

FEATURES
source
    cdna library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: National Institutes of Health Intramural
    Sequencing Center (NISC),
    Gaithersburg, Maryland;
    Web site: http://www.nisc.nih.gov/
    Contact: nisc.mgc@nih.gov
    Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
    Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
    Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
    Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
    McDowell, J., Pearson, R., Snyder, B., Stantrilop, S., Thomas, P.J.,
    Tjongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,
    Zhang, L., H. and Green, E.D.

    Clone distribution: MGC clone distribution information can be found
    through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
    Series: IRAL Plate: 1 Row: 9 Column: 13
    This clone was selected for full length sequencing because it
    passed the following selection criteria: matched mRNA gi: 1905802.
    Location/Qualifiers
        1..2031
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /db_xref="IMAGE:2821792"
        /clone="MGC:8478 IMAGE:2821792"
        /tissue-type="lung, small cell carcinoma"
        /clone_id="NIH-MGC-7"
        /lab_host="DH10B-R"
        /note="Vector: pOTB7"
        139..1281
        /codon_start=1
        /product="flap structure-specific endonuclease 1"
        /protein_id="AAH00323.1"
        /db_xref="GI:12653113"
        /translation="MTGQGLAKIADVAASAIKRENDIKYRGRKVAIDASLKIOTLI
        AVRGQVLDNERETTSILMGAFRTIKRMKEIKRPYVFGKPOLKSELKRSR
        RRAEAKOQLOAOAGAEDVEKFTRLVKTRONHDEKRLSLAMGIPYDASEAB
        ASCEALVKAKVYAATVEDMDCTFGSPVLMHRLVASEAKLPIQEFILSLDELG
        NOEFDVLCILLSDYCESIRGIGPRAVDLQKHRSLEIVRLDPKRYVPEVNH
        KEAHLFLEPEVDELPSEVLMKSEPNBEELIKFMCGEKFSEIRSRISRKLSRQ
        QKRALDPFQVTVSSLSAKRKEPEKSGSKKKAKTGAAGKFRKRG"

BASE COUNT      533 a      465 c      589 g      444 t
ORIGIN
Query Match      20.7%      Score 302.8;      DB 9;      Length 2031;
Best Local Similarity 55.0%;      Pred. No. 4,7e-64;
Matches 661;      Conservative 0;      Mismatches 532;      Indels 9;      Gaps 3;
OY 85 ATGGGATCAAGGTTTACGAAACTGCTGCGGACATGCGCCCAAGGCGATGAAGAG 144
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 ATGGGATCAAGGCTGCGCAACTATATCTGATGTGGCCCAAGTCCATCGGAGAG 198
OY 145 CAGAGTTGAGAGCTACTTGGCGCGCAAAATCGCGTGCAGCGCAGATGACATATAC 204
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 AATGATATCAAGAGCTCTTGGCGGTAAGGTGGCAATGATGCTCTATGACCATTTAT 258
OY 205 CAGTTCTGATGTGTGGAAGAGCAGTGAACACTCTCAACAAATGAAGCTGTGTGA 264
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 CAGTTCTGATGTGTGGAAGAGCAGTGAACACTCTCAACAAATGAAGCTGTGTGA 315
OY 265 GTCATGATGATTTGCAAGGATGTCAACCGGACATTAAGATTAAGTGAAGGCGAATC 324
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 ACCACGACGACCTGATGGAGATGTTTACCGCACATGCTGATGATGAGAGCGCATC 375
OY 325 AAGCCAGTTATTTTATGATGGCAAGGCTCTGATATGAAGAAACAAGACTTGCTA 384
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 AAGCCGCTATGCTTTGATGAGCAAGCGCCACAGCTCAAGTACGAGCGAGCTGCGCAA 435
OY 385 ACATATCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGTAGAGTAGAGAT 444
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 CGCAGTGAAGCGGCGGCTGAGGCGAGAGACGCTGACAGAGCTCAGGCTGCTGGGGCC 495
OY 445 AAGATGCGATGAAAAATTTAGCAAGAGAGACTGTAAAGTCAACAAGCAACAAGAA 504

```


* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

```
* 1 20726: contig of 20726 bp in length
* 20727 20826: gap of 100 bp
* 20827 34666: contig of 13840 bp in length
* 34667 34766: gap of 100 bp
* 34767 46065: contig of 11299 bp in length
* 46066 46165: gap of 100 bp
* 46166 53516: contig of 7351 bp in length
* 53517 53616: gap of 100 bp
* 53617 61065: contig of 7449 bp in length
* 61066 61165: gap of 100 bp
* 61166 67998: contig of 6833 bp in length
* 67999 68098: gap of 100 bp
* 68099 74526: contig of 6428 bp in length
* 74527 74626: gap of 100 bp
* 74627 81299: contig of 6673 bp in length
* 81300 81399: gap of 100 bp
* 81400 87316: contig of 5917 bp in length
* 87317 87416: gap of 100 bp
* 87417 91541: contig of 4125 bp in length
* 91542 91641: gap of 100 bp
* 91642 95731: contig of 4090 bp in length
* 95732 95831: gap of 100 bp
* 95832 101128: contig of 5297 bp in length
* 101129 101228: gap of 100 bp
* 101229 104948: contig of 3720 bp in length
* 104949 105048: gap of 100 bp
* 105049 110306: contig of 5258 bp in length
* 110307 110406: gap of 100 bp
* 110407 112969: contig of 2563 bp in length
* 112970 113068: gap of 100 bp
* 113070 116484: contig of 3415 bp in length
* 116485 116584: gap of 100 bp
* 116585 119599: contig of 3015 bp in length
* 119600 119699: gap of 100 bp
* 119700 121881: contig of 2182 bp in length
* 121882 121981: gap of 100 bp
* 121982 122703: contig of 722 bp in length
* 122704 122803: gap of 100 bp
* 122804 125566: contig of 2763 bp in length
* 125567 125666: gap of 100 bp
* 125667 126845: contig of 1179 bp in length
* 126846 126945: gap of 100 bp
* 126946 129349: contig of 2404 bp in length
* 129350 129449: gap of 100 bp
* 129450 131268: contig of 1819 bp in length
* 131269 131368: gap of 100 bp
* 131369 132565: contig of 1197 bp in length
* 132566 132665: gap of 100 bp
* 132666 134031: contig of 1366 bp in length
* 134032 134131: gap of 100 bp
* 134132 136176: contig of 2045 bp in length
* 136177 136276: gap of 100 bp
* 136277 137418: contig of 1142 bp in length
* 137419 137518: gap of 100 bp
* 137519 138627: contig of 1109 bp in length.
```

FEATURES

source

```
1..138627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="GMB9-22K5"
misc_feature
1..20726
/note="assembly_fragment"
misc_feature
20827..34666
/note="assembly_fragment"
misc_feature
34767..46065
/note="assembly_fragment"
```

```
misc_feature 46166..53516
/note="assembly_fragment"
misc_feature 53617..61065
/note="assembly_fragment"
misc_feature 61166..67998
/note="assembly_fragment"
misc_feature 68099..74526
/note="assembly_fragment"
misc_feature 74627..81299
/note="assembly_fragment"
misc_feature 81400..87316
/note="assembly_fragment"
misc_feature 87417..91541
/note="assembly_fragment"
misc_feature 91642..95731
/note="assembly_fragment"
misc_feature 95832..101128
/note="assembly_fragment"
misc_feature 101229..104948
/note="assembly_fragment"
misc_feature 105049..110306
/note="assembly_fragment"
misc_feature 110407..112969
/note="assembly_fragment"
misc_feature 113070..116484
/note="assembly_fragment"
misc_feature 116585..119599
/note="assembly_fragment"
misc_feature 119700..121881
/note="assembly_fragment"
misc_feature 121982..122703
/note="assembly_fragment clone_end:17 vector_side:right"
misc_feature 122804..125566
/note="assembly_fragment"
misc_feature 125667..126845
/note="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature 126946..129349
/note="assembly_fragment"
misc_feature 129450..131268
/note="assembly_fragment"
misc_feature 131369..132565
/note="assembly_fragment"
misc_feature 132666..134031
/note="assembly_fragment"
```

Query Match 20.7%; Score 302.8; DB 2; Length 138627;
Best Local Similarity 55.0%; Pred. No. 1.2e-63;
Matches 661; Conservative 0; Mismatches 532; Indels 9; Gaps 3;

```
QY 85 ATGGGATCAAGGCTTTGACGAACCTGCTGGCGGCAATGCGCCCAAGGCGATGAAGGAG 144
DB 17322 ATGGGATTCAGAGCGCTCGCCCAACTATTCCTATGTGGCCCCCAGTCCGCGGAG 17381
QY 145 CAGAAATTCGAGAGCTACTCGCGCGCAAAATGCGCCGTCAGCGGACATGACCATATAC 204
DB 17382 AATGACATCAAGAGCTACTTTGGCGGTAGGTGGCGCATTCGATGCTCTATGACCATTTAT 17441
QY 205 CAGTTCCTGATGTGTGTGGGAAGCAGGACATGGAAACTGCACAAATGAAGGTGGTAA 264
DB 17442 CAGTTCCTGATGTGTGTGGGAAGCAGGACATGGAAACTGCACAAATGAAGGTGGTAA 264
QY 17442 CAGTTCCTGATGTGTGTGGGAAGCAGGACATGGAAACTGCACAAATGAAGGTGGTAA 264
QY 265 GTGACTAGTCAATTCGACGAAGATGTTCAACCGGACAAATAGATTTAGTGAAGCGGGAATC 324
DB 17499 ACCACAGCGACACTGATGGGATGCTATTCACCGCACATTCGATGATGGAAGCGGATC 17558
QY 325 AAGCCAGTTTATGTTTGTGATGGCAAGCGCTCGATATGGAAGAAACAAGAGCTTGTAA 384
DB 17559 AAGCCAGTTTATGTTTGTGATGGCAAGCGCTCGATATGGAAGAAACAAGAGCTTGTAA 384
QY 385 AGATACCTCAAAAAGATGATGACACCAAGATCTGACTGAGCGCATGATGAGTAGAGAT 444
DB 17619 CGAGATGACCGCGGCGGAGGACAGAGAGCTGACGACGAGCTGAGCTGTGGGGCC 17678
QY 445 AAGATGATGATGAAATAATTGACGAAGAGACTGTAAAGTCAACAGGCAACACAGAA 504
```

Db 17679 GAGCAGAGGTGGAAATTTCACTAAGCGGCTGGTGAAGGTACTAAGCAGCATATAT 17738
QY 505 GATTGTAACGGCTATTAGACTTATGSGGGTCTCTGTTAGAGGACTTTGTAAGCA 564
Db 17739 GAGGCGAAATCTGCTGAGCCCTCATGGCAATCCCTATCTGATGACACCCAGTGAGCA 17798
QY 565 GAGCAGCAATGTGAGCGCTTTGCATTAACGATTAAGGTGTGCGCTGCTTCAAGAT 624
Db 17799 GAGGCGAGCTGTGCTGCGCTGGTGAAGGCTGGCAAGGTATGCTGCGGCTACCGAGAC 17658
QY 625 ATGAGCTCCTTACTTTTGGGCTCCAGGCTTCTTCTGCTATTAATGATCAAGTTCC 684
Db 17859 ATGAGCTCCTTACTTTGCGGACGCTTGTGCTATTAATGCGACCTGCTCCAGTAAAC 17918
QY 685 AAGAAATACCTGTGATGATTTGCAATTTGCAAGTTTGGAGAGCTTGAATCTACC 744
Db 17919 AAAAGCTGCATTCAGAGATTCACCTGACCGGATTTCTGAGAGCTGGGCTGAC 17978
QY 745 ATGAGCAGTTTCAATGATTTTGTGATCTGTGTGATGATTTGATGATCAATCAA 804
Db 17979 CAGGACAGATTTGATGCTGATCTGATCTGATGAGCAGCTGATGATGATCCGG 18038
QY 805 GGTATCGGGGCGCAACAGCTGTAATCTTATCTGATGATGCTGATGATGATGATGAT 864
Db 18039 GGTATTTGGGCGCAAGCGGCTGTGACCTCATCCAGACAGACAGATGAGAGATC 18098
QY 865 TTGAGATCTTAAATGAAGACAGATATCAATTTCTGAGAGCTGGCTTACCAAGACT 924
Db 18099 GTGGGGGCTTGAACCCAGATGATGATGATGATGATGATGATGATGATGATGAT 18158
QY 925 CGAGCTTTGTTCAAGGAGCTTATG--TCACATTTGATTTCTGATGATGATGATGAT 981
Db 18159 CACAGCTCTTCTTGTGAGCTGAGTGTGAGCCGAGCTCTGTGAGCTGAGTGAAG 18218
QY 982 GCACCTATGAGAGGCTCTCATTAAGTTCTGTTAAAGATTAATGTTTACAGCAAT 1041
Db 18219 GAGCAATATGAAGAGCTGATCAATGATGATGATGATGATGATGATGATGATGAT 18278
QY 1042 CGGCTGCAAGGCGCATAGAGATCAATCTGCCAAGATTAATCTGCGCAAGAGA 1101
Db 18279 CGATTCGCGAGTGGGCTGAGAGCTGAGTGAAGAGCGCGCAAGGAGCGCGCG 18338
QY 1102 CTGAGAGCTTTTCAAGCACTGCCAC--CAGATGAGCAGCGCTAAAGAGAGAG 1158
Db 18339 CTGATGATTTCTTCAAGGCTGAGCTGATCTCTTCAAGGAGAGAGAGAGAGAG 18398
QY 1159 ACTTCGATTAACAGAGGAGCTGCGAACAAGAAACAGAGAGCTGTTGAAGAG 1218
Db 18399 CCCAAGGATCCACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 18458
QY 1219 AATTAATCTTGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1278
Db 18459 AATTAATGTTGTTCCCATTAATCTCTTCAAGGAGAGAGAGAGAGAGAGAGAG 18518
QY 1279 TT 1280
Db 18519 TT 18520

RESULT 15
AP002380/c 169053 bp DNA linear HTG 31-MAY-2000
LOCUS Homo sapiens chromosome 11 clone RP11-467L20 map 11q, WORKING DRAFT
DEFINITION SHORUCE, 35 unordered pieces.
ACCESSION AP002380
VERSION AP002380.1 GI:8131644
KEYWORDS HTG, HTGS_PHASE1, HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-467L20.
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 169053)

AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (29-May-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail: hattori@gsc.riken.go.jp,
URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDrafl1
Center clone name: RP11-467L20
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator EM-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 147827 bases at least Q40
Consensus quality: 157936 bases at least Q30
Consensus quality: 162410 bases at least Q20
Insert size: 165653; sum-of-contigs
Quality coverage: 4.43x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
35 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1
21322 contig of 21322 bp in length
40245 contig of 18823 bp in length
40346 contig of 10147 bp in length
50492 contig of 7977 bp in length
58569 contig of 8716 bp in length
58670 contig of 7753 bp in length
67486 contig of 9105 bp in length
75349 contig of 8453 bp in length
84554 contig of 6833 bp in length
89593 contig of 6086 bp in length
96546 contig of 4468 bp in length
102631 contig of 4521 bp in length
102732 contig of 5148 bp in length
107300 contig of 3733 bp in length
111820 contig of 3071 bp in length
111921 contig of 3439 bp in length
117169 contig of 12761 bp in length
120901 contig of 12940 bp in length
121002 contig of 133141 bp in length
124173 contig of 133242 bp in length
127712 contig of 136380 bp in length
129502 contig of 2945 bp in length
133242 contig of 3756 bp in length
136481 contig of 3068 bp in length
139425 contig of 2279 bp in length
14281 contig of 2117 bp in length
143582 contig of 1563 bp in length
146449 contig of 1166 bp in length
146550 contig of 1524 bp in length
148929 contig of 15473 bp in length
151146 contig of 15496 bp in length
151416 contig of 15673 bp in length
153372 contig of 2269 bp in length
153473 contig of 2117 bp in length
15496 contig of 1563 bp in length
155097 contig of 1166 bp in length
156763 contig of 1524 bp in length
159132 contig of 15473 bp in length
159233 contig of 15496 bp in length
161349 contig of 2279 bp in length
161450 contig of 2117 bp in length
162951 contig of 1563 bp in length
164634 contig of 1166 bp in length
165829 contig of 1201 bp in length
165930 contig of 1060 bp in length
167130 contig of 1060 bp in length
168290 contig of 1060 bp in length

Db 3399 GACGACGAGGTGGAAAAATTAATCACTAAGCGGCTGTGAAGGTCTACTAAGCAGACACATATGAT 3340
QY 505 GATTGTAAACGGCTATTAAAGACTTATGGGGGTTCCCTGTTAGAGGACCTTCTGAAGCA 564
Db 3339 GAGTGCAAACATCTGTGAGCCCTCATGGGCATCCCTTATCTGATGACACCCAGTAGGCA 3280
QY 565 GAAGCAGAATGTGCACCCCTTTCATTAACGATAAGGTGTTCGCTGTGCTTCAGAAAT 624
Db 3379 GAGGCCAGCTGTGCTGCTGTGAAGGCTGGCAAGTCTATGCTGCGGCTCCAGAGAC 3220
QY 625 ATGAGCTCCCTTACTTTTGGGGCTCCACGGTCTTCCTGCTATTAATGATCCAAGTTC 684
Db 3219 ATGAGCTGCCCTCACTTCGAGCGCCCTGTCTAATGCGACACCTGACCTGCCAGTAAAGCC 3160
QY 685 AAGAAATACCTGTGATGATTTGATGTTGCCAAGTTTGGAGGAGCTGAACCTCAC 744
Db 3159 AAAAAGCTGCAATCCAGGATTCACCTGAGCGGATTCCTCAGAGACCTGGCCCTGAC 3100
QY 745 ATGACACAGTTTCATGATTTGTGCATCCTGTGTGATGTGACTATTTGATAGCATCAA 804
Db 3099 CAGGAACAGTTTGTGATCTGTGCATCCTGTAGCAGAGTACTGTAGAGTATCCGG 3040
QY 805 GSTATGGGGGGCCAAACGCTGAAACTTATTCGTAACATGGTCCATAGAAAAGCATC 864
Db 3039 GGTATTGGGCCCAAGGGGCTGTGGACCTCATCCAGAAAGCACAAGAGCATGAGAGATC 2980
QY 865 TTGGAGAACTTATTAAGAAGATATCAATTCCTGAGGAGCTGGCTTACCAGAAAGCT 924
Db 2979 GTGGGGGCACTTGACCCCAACAGTACCTGTGCCAGAAATTTGGCTCCACAAGAGGCT 2920
QY 925 CGAGGCTTTCTCAAGGAGCCTAATG--TCACATTGATATTCCTGAGCTAAATGACT 981
Db 2919 CACCAGCTCTTCTTGAGACTGAGTGTGAGCCAGAGACTGTGTGAGCTGAAGTGGAGC 2860
QY 982 GCACCTGATGAGGAGGCTCATAGTTCTGTGTAAGATATATGTTCAACGAGAT 1041
Db 2859 GAGCCAATGATGAAGAGAGCTGATCAATGATGTGTGAAGAGCAGTTCCTGAGAG 2800
QY 1042 CGGCTGACAAAGGCCATAGAGATCAATTCGCCAAGATTAATCGTCCAGAGAGA 1101
Db 2799 CGAATCCGCGAGTGGGTCAAGAGCTGAGTAAGAGCCGCCAAGGACGACCCAGGGCCG 2740
QY 1102 CTGAGTCTTTTTCAGGCCAAGTCCAC--CACATCAGACCGCTAAAGCGAAGAG 1158
Db 2739 CTGATGATTTCTTCAAGGTGACCGGCTCCTCTTCAGTAAAGCGCAAGAGCCAGAA 2680
QY 1159 ACTTCGATTAACAGCAGGAGGAGTGCAGCAAGAAACAAAGGCTGTGGAAGAG 1218
Db 2679 CCGAAGGAGATCCACTAAGAAAGAGACTGGGGCGACGAGGAAGTTTAAAGGGGA 2620
QY 1219 AAATAATCTTGATGCTTGATGACACTAGACTAGCAAGCAGCGGTGGCGTGTATCAG 1278
Db 2619 AAATAATGTTTCCCATTAATCCCTTACCCAGAGATATTTGCCGTTGTACCC 2560
QY 1279 TT 1280
Db 2559 TT 2558

Search completed: November 5, 2002, 12:45:26
Job time : 2515.65 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:10:48 : Search time 174.666 Seconds
(without alignments)
14380.822 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463
Sequence: 1 caccagaaatgctcgcgcgc.....aaaaaaaaaaaaaaaaaaaa 1463

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_032802: *
1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1463	100.0	1463	21	Maize Rad2/FEN-1 c
2	1421.8	97.2	1478	21	Maize Rad2/FEN-1 c
3	1409.8	96.4	1541	21	Maize Rad2/FEN-1 c
4	1350.2	92.3	1381	21	Maize Rad2/FEN-1 c
5	302.6	20.7	1144	20	Human FEN-1 DNA fr
6	302.6	20.7	1144	20	Human FEN-1 CDNA
7	276.6	18.9	2033	20	Human FEN-1 genom
8	276	18.9	1300	23	Drosophila melanog
9	272.2	18.6	1930	20	Mouse FEN-1 CDNA

10	270.2	18.5	1149	20	AAK02109
11	252.6	17.3	3362	23	ABL14286
12	173.2	11.8	5471	23	ABL10094
13	159.0	10.9	1023	18	AAAT76685
14	159.6	10.9	1023	18	AAAT76685
15	159.6	10.9	1023	18	AAAT76685
16	159.6	10.9	1023	18	AAAT76685
17	149.6	10.2	1115	19	AAV53984
18	149.6	10.2	1115	19	AAV53984
19	139.6	9.5	510	22	AAK37479
20	135.6	9.3	1032	20	AAI18286
21	134	9.2	34980	22	AAH41224
22	128	8.7	514	19	AAV53990
23	126.4	8.6	1164	19	AAV53975
24	125.6	8.6	1115	19	AAV53978
25	117.6	8.0	1729	23	AAK3732
26	110.2	7.5	889	19	AAV53974
27	98.6	6.7	1011	19	AAV5913
28	97.6	6.7	1011	19	AAV54009
29	97.6	6.7	1164	19	AAV53989
30	94.8	6.5	386	19	AAV53979
31	90	6.2	546	22	ABA6365
32	90	6.2	546	22	AAK1490
33	90	6.2	546	22	AAK37269
34	90	6.2	546	22	AAI18098
35	89.2	6.1	981	18	AAAT76682
36	89.2	6.1	981	18	AAV53948
37	86.6	5.9	243	22	AAK50354
38	86.6	5.9	243	22	AAI27416
39	86	5.9	981	19	AAV5837
40	76.2	5.2	777	19	AAV54010
41	74.6	5.1	840	19	AAV53977
42	69.4	4.7	987	19	AAV54011
43	67.8	4.6	296	19	AAV53976
44	62.8	4.3	514	19	AAV53985
45	62.4	4.3	4551	21	AAV70129

ALIGNMENTS

RESULT 1	AAA27923	standard: CDNA; 1463 BP.
ID	AAA27923	
XX	AAA27923	
AC	AAA27923	
XX	AAA27923	
DT	12-SEP-2000	(first entry)
XX	12-SEP-2000	
DE	Maize Rad2/FEN-1 CDNA.	
XX	Maize Rad2/FEN-1	
KW	Maize; Rad2/FEN-1; transgenic plant; male sterile plant;	
RW	endonuclease; exonuclease; DNA repair; gene targeting; ss.	
XX		
OS	Zea mays.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	85..1224
XX		/*tag= a
PN	WO200036109-A1.	
XX		
PD	22-JUN-2000.	
XX		
PF	16-NOV-1999;	99WO-US27147.
XX		
PR	15-DEC-1998;	98US-0112332.
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	MahaJan FB;	
XX		
DR	WPI; 2000-452026/39.	

Yeast FEN-1 CDNA.
Drosophila melanog
Drosophila melanog
Pyrococcus furiosu
Pyrococcus furiosu
Nucleotide sequenc
P. furiosus N-term
Human bone marrow
Probe #8219 for ge
Pyrococcus heat re
Pyrococcus abyssi
Nucleotide sequenc
Nucleotide sequenc
P. furiosus N-term
DNA encoding novel
Nucleotide sequenc
Archaeoglobus fulg
Nucleotide sequenc
M. jannaschii N-te
Nucleotide sequenc
Human fetal liver
Human bone marrow
Probe #8031 for ge
Nucleotide sequenc
Human bone marrow
Probe #17349 for g
Methanococcus jann
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Plasmodium falcipar

DR P-PSDB; AAY95307.
XX Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX
PS Claim 1; Page 69-71; 85pp; English.
XX
CC The present sequence is that of maize cDNA coding for RAD2/FEN-1
CC (see AAY95307). The corresponding RNA was isolated from immature
CC ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a
CC structure specific endonuclease which under certain conditions also
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC joining. The protein is also useful in strand exchange
CC reactions involving homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;
Query Match 100.0%; Score 1463; DB 21; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGGAGATAGTCTCGGCGCGGCTTCTTGCGCAGTCCGGCTCAGCCCGCCGCCA 60
DB 1 CAGGAGATAGTCTCGGCGCGGCTTCTTGCGCAGTCCGGCTCAGCCCGCCGCCA 60
QY 61 CCCGCCACGCGCGCCGACACGAGATGGGATCAAGGGTTAGCAAACTGTGCGGAC 120
DB 61 CCCGCCACGCGCGCCGACACGAGATGGGATCAAGGGTTAGCAAACTGTGCGGAC 120
QY 121 AATGCGCCCAAGGCGATGAGAGGAGCAGAACTTGCAGAGCTTCTGCGCGCAATTCGCC 180
DB 121 AATGCGCCCAAGGCGATGAGAGGAGCAGAACTTGCAGAGCTTCTGCGCGCAATTCGCC 180
QY 181 GTGCGACGACGATGAGCATATACAGTTCCTGATTTGTGTTGGAAGGACAGCATGGA 240
DB 181 GTGCGACGACGATGAGCATATACAGTTCCTGATTTGTGTTGGAAGGACAGCATGGA 240
QY 241 ACTCTCAAAATGAGCTGTGTAAGTCACTAGTTCATTTCAGGAATGTTCAACGAGACA 300
DB 241 ACTCTCAAAATGAGCTGTGTAAGTCACTAGTTCATTTCAGGAATGTTCAACGAGACA 300
QY 301 ATAGATTACTGAGACGCGGAATCAAGCAGTTATGTTTGGTGAAGCCTCTGAT 360
DB 301 ATAGATTACTGAGACGCGGAATCAAGCAGTTATGTTTGGTGAAGCCTCTGAT 360
QY 361 ATGAGAAAGAAAGAGCTGTGTAAGATCTCAAAAAGAGATGATCAACCAAGATCTG 420
DB 361 ATGAGAAAGAAAGAGCTGTGTAAGATCTCAAAAAGAGATGATCAACCAAGATCTG 420
QY 421 ACTGAGCAGTATGAGTATGAGATTAAGATGATGAAATTTGAGCAAGAGACTGTA 480
DB 421 ACTGAGCAGTATGAGTATGAGATTAAGATGATGAAATTTGAGCAAGAGACTGTA 480
QY 481 AAGGTCAAGAGCAACACGAAGATTGAAACGCGTATTAGACTTATGGGGTTCCT 540
DB 481 AAGGTCAAGAGCAACACGAAGATTGAAACGCGTATTAGACTTATGGGGTTCCT 540
QY 541 GTTGTATAGGACCTTCTGAGAGCAGAGCAAGATGTGACCCCTTGACATTAACGATTA 600
DB 541 GTTGTATAGGACCTTCTGAGAGCAGAGCAAGATGTGACCCCTTGACATTAACGATTA 600

QY 601 GTTTCGCTGTGCTTCTTCAGAGATATGAGTCCCTTACTTTTGGGCGTCCAGGTTCTT 660
DB 601 GTTTCGCTGTGCTTCTTCAGAGATATGAGTCCCTTACTTTTGGGCGTCCAGGTTCTT 660
QY 661 CGTCATTATGATGATCCAAAGTTCGAAGAAATACCTGTATGGAATTTGATGTCACAG 720
DB 661 CGTCATTATGATGATCCAAAGTTCGAAGAAATACCTGTATGGAATTTGATGTCACAG 720
QY 721 GTTTTGGAGAGCTTGAGTCACTGACGACGCTTCATTTGATTTGATCCTGTGCGA 780
DB 721 GTTTTGGAGAGCTTGAGTCACTGACGACGCTTCATTTGATTTGATCCTGTGCGA 780
QY 781 TGTGACTATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 TGTGACTATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 CAACATGGGTCATAGAAAGCATTTGGAATCTTAAATTAAGACAGATATCAATTTCT 900
DB 841 CAACATGGGTCATAGAAAGCATTTGGAATCTTAAATTAAGACAGATATCAATTTCT 900
QY 901 GAGGACTGGCTTACCAAGAGCTGAGGCTGTTCAAGAGCTTAATGTCATTTGAT 960
DB 901 GAGGACTGGCTTACCAAGAGCTGAGGCTGTTCAAGAGCTTAATGTCATTTGAT 960
QY 961 ATTCTGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 ATTCTGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 GATTAAGTCTTCAACGAGATGAGGCTGACAAAGCCATGAGAGATCAATCTGCCAAG 1080
DB 1021 GATTAAGTCTTCAACGAGATGAGGCTGACAAAGCCATGAGAGATCAATCTGCCAAG 1080
QY 1081 AATTAATCTGTCGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 AATTAATCTGTCGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 CCGCTAAACGGAAGGAGTCTGGATTAACAAAGCAGACGCTGCAACAAAGAAACA 1200
DB 1141 CCGCTAAACGGAAGGAGTCTGGATTAACAAAGCAGACGCTGCAACAAAGAAACA 1200
QY 1201 AAGCTGTGGAAGAGAAATATATCTTGATGATGATGATGATGATGATGATGAT 1260
DB 1201 AAGCTGTGGAAGAGAAATATATCTTGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CAGCGTGGCGTATCACTTCTGATTAATTAATTAATTAATTAATTAATTAATTA 1320
DB 1261 CAGCGTGGCGTATCACTTCTGATTAATTAATTAATTAATTAATTAATTAATTA 1320
QY 1321 GGTAAAGTTTGCATGTTTCAAGCTGGGTAAGTATGTTGTTGAAGAGATTGCTG 1380
DB 1321 GGTAAAGTTTGCATGTTTCAAGCTGGGTAAGTATGTTGTTGAAGAGATTGCTG 1380
QY 1381 TACCAAGTAACAAACTATGCTGTTTAACTTCTGTTGAAGTAAAGAAAAA 1440
DB 1381 TACCAAGTAACAAACTATGCTGTTTAACTTCTGTTGAAGTAAAGAAAAA 1440
QY 1441 AAAAAA 1463
DB 1441 AAAAAA 1463
RESULT 2
AAA27926
ID AAA27926 standard; cDNA: 1478 BP.
XX AAA27926;
AC
XX 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 cDNA.
XX
XX Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.

XX Zea mays.

XX Key Location/Qualifiers

FH CDS 97..1236

FT /tag= a

XX WO200036109-A1.

XX 22-JUN-2000.

XX 16-NOV-1999: 99WO-US27147.

XX 15-DEC-1998: 98US-0112332.

XX (PION-) PIONEER HI-BRED INT INC.

XX Mahajan PB:

XX WPI: 2000-452026/39.

XX P-PSDB: AAY95310.

XX The present sequence is that of maize cDNA coding for RAD2/FEN-1 (see AAY95310). The cDNA was isolated from a library prepared from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a structure specific endonuclease which under certain conditions also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to produce the Rad2/FEN-1 polypeptides in transgenic plant cells. The protein is involved in the regulation of DNA repair and recombination in plant systems and therefore may be used for improving gene targeting during further recombinant DNA protocols in DNA replication and nucleotide excision and repair reactions. CC The exolytic activity is involved in double strand break repair and end joining. The protein is also useful in strand exchange reactions during homologous recombination. These functions may be useful in gene targeting and in the production of male sterile plants. CC The efficacy of gene targeting can be improved by the overexpression of exogenous Rad2/FEN-1 while male sterile plants can be produced by the down regulation of Rad2/FEN-1 expression.

XX Example 1: Page 77-79; 85pp; English.

Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;

Query Match 97.2%; Score 1421.8; DB 21; Length 1478;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 AATAGCTCGGGGGGTTTCTTGGCCATCCGGCTAGCCGCCGCCGACCCGCC 66

Db 19 AATAGCTCGGGGGTGGCTTTCTTGGCCATCCGGCTAGCCGCCGCCGACCCGCC 78

QY 67 ACAAGCCGCCAGACGATGCGATCAAGGGTTTGACGAAACTCTGCGGACATCGC 126

Db 79 ACAGCCGCCAGACGATGCGATCAAGGGTTTGACGAAACTCTGCGGACATCGC 138

QY 127 CCCAAGCGGATGGAAGGACGAGTCTGAGACTACTTGGCCGCAAAATCGCGTGCAC 186

Db 139 CCCAAGCGGATGGAAGGACGAGTCTGAGACTACTTGGCCGCAAAATCGCGTGCAC 198

QY 187 GCCAGCATGACATATACAGTTCCTGATTTGATTTGAAGACAGGATGAAACTCTC 246

Db 199 GCCAGCATGACATATACAGTTCCTGATTTGATTTGAAGACAGGATGAAACTCTC 258

QY 247 ACAATGACCTGGTGAAGCACTACTGCAATTTGCAAGAAATGTAACCGGACATAGA 306

Db 259 ACAATGACCTGGTGAAGCACTACTGCAATTTGCAAGAAATGTAACCGGACATAGA 318

QY 307 TTAGTGAAGCGGGATATCAAGCCAGTTTATGTTTGGATGGCAAGCCTCCTGATATGAAG 366

Db 319 TTAGTGAAGCGGGATATCAAGCCAGTTTATGTTTGGATGGCAAGCCTCCTGATATGAAG 378

QY 367 AACAGAGCTTGGCTAAAGATGACTCAAAAAAGAGATGATGCAACCAAGATGCTGACTGAG 426

Db 379 AACAGAGACTTGGCTAAAGATGACTCAAAAAAGAGATGATGCAACCAAGATGCTGACTGAG 438

QY 427 GCAGTAGAGGTAGAGATTAAGATGCGATTGAAAATTTGACGAAGGACTGTAAAGTTC 486

Db 439 GCAGTAGAGGTAGAGATTAAGATGCGATTGAAAATTTGACGAAGGACTGTAAAGTTC 498

QY 487 ACAAGCAACACACGAGGATTTGTAAGCGCTATTAAGACTTATGAGGGGCTCTGTTGTA 546

Db 499 ACAAGCAACACACGAGGATTTGTAAGCGCTATTAAGACTTATGAGGGGCTCTGTTGTA 558

QY 547 GAGGCACTTTCGAAAGCAAGCAAGATGTGACGCCCTTTCGATTAACGATTAAGTGTTC 606

Db 559 GAGGCACTTTCGAAAGCAAGCAAGATGTGACGCCCTTTCGATTAACGATTAAGTGTTC 618

QY 607 GCTGTGCTTCGAAAGATATGACTCCCTTACTTTTGGGCTCCAGGTTCTTCTGTCAT 666

Db 619 GCTGTGCTTCGAAAGATATGACTCCCTTACTTTTGGGCTCCAGGTTCTTCTGTCAT 678

QY 667 TTAATGATCCCAAGTTCACAAAGAAATACCTGTGATGTAATTTATGTGCAAGGTTTG 726

Db 679 TTAATGATCCCAAGTTCACAAAGAAATACCTGTGATGTAATTTATGTGCAAGGTTTG 738

QY 727 GAGGAGCTTGAACCTACCATGAGACAGTTCAATTTTGTGATCCGTTGTGATGTGAC 786

Db 739 GAGGAGCTTGAACCTACCATGAGACAGTTCAATTTTGTGATCCGTTGTGATGTGAC 798

QY 787 TATTGTGATGATCAATGAAGGTATCGGGGGCAACACTCTGAAACTTATTCCTCAACAT 846

Db 799 TATTGTGATGATCAATGAAGGTATCGGGGGCAACACTCTGAAACTTATTCCTCAACAT 858

QY 847 GGCTCCATGAAGACATTTTGAGATCTTAATTAAGACAGATATCAAAATTCCTGAGGAC 906

Db 859 GGCTCCATGAAGACATTTTGAGATCTTAATTAAGACAGATATCAAAATTCCTGAGGAC 918

QY 907 TGGCCTTACCAAGAGCTCGACGCTTGTTCAGGAGGCTATGTCACATTTGATTTCTCT 966

Db 919 TGGCCTTACCAAGAGCTCGACGCTTGTTCAGGAGGCTATGTCACATTTGATTTCTCT 978

QY 967 GAGCTAAATGAGACTGCACTGATGAGAGGGTCTCAATAGTTTCTGTTAAAGATTAAT 1026

Db 979 GAGCTAAATGAGACTGCACTGATGAGAGGGTCTCAATAGTTTCTGTTAAAGATTAAT 1038

QY 1027 GGTTCACAGGAATCGGGTGAACAAGGCAATGAGAGAAATCAATCTGCCAGATATAA 1086

Db 1039 GGTTCACAGGAATCGGGTGAACAAGGCAATGAGAGAAATCAATCTGCCAGATATAA 1098

QY 1087 TCGTGCAGAGGAAGCTGAGTCTTTTTCAGGCCAATGCGCACATATAGCACCGCTTA 1146

Db 1099 TCGTGCAGAGGAAGCTGAGTCTTTTTCAGGCCAATGCGCACATATAGCACCGCTTA 1158

QY 1147 AAAGGAGAGACACTTCGATTAACAAGCAAGGCAAGCTCGAACAAGAAACAAAGGCT 1206

Db 1159 AAAGGAGAGACACTTCGATTAACAAGCAAGGCAAGCTCGAACAAGAAACAAAGGCT 1218

QY 1207 GGTGGAAGAAAGAAATATCTTGGATGCTTATGATGATGATGATGATGATGATGATGAT 1266

Db 1219 GGTGGAAGAAAGAAATATCTTGGATGCTTATGATGATGATGATGATGATGATGATGAT 1278

QY 1267 TGGCGTGAATCTGCTTAAGATTAATTAACCTGCTTTTAACCTGAGCTTGTGTAA 1326

Db 1279 TGGCGTGAATCTGCTTAAGATTAATTAACCTGCTTTTAACCTGAGCTTGTGTAA 1338

QY 1327 AGTTGCTGATGTTCAAGCTGGGGTAGTAAATTTGTTGTTGAAGATGTTGTTGTTGTTG 1386

Db 1339 AGTTGCTGATGTTCAAGCTGGGGTAGTAAATTTGTTGTTGAAGATGTTGTTGTTGTTG 1398

QY 1387 GTACAAACCTTATGCTGTGTTTCTTCTTGTCTTGAAGTAAAAAATAAAAA 1446

```
|||||
Db 1399 GTAAACAACCTTATCGCTGTTTCTTCTGCTTGAAGTATGTATGCCAGATAAAA 1458
QY 1447 AAAAAAAAAAAAAAAAAA 1463
    |||||||
Db 1459 AAAAAAAAAAAAAAAAAA 1475

RESULT 3
AAA27924
ID AAA27924 standard; cDNA; 1541 BP.
AC AAA27924;
AT 12-SEP-2000 (first entry)
DE Maize Rad2/FEN-1 cDNA.
XX
XX Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KM endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX
XX Zea mays.
OS
XX Key Location/Qualifiers
FH 79.1218
FT CDS /*tag= a
XX
XX WO200036109-A1.
XX
XX 22-JUN-2000.
XX
XX 16-NOV-1999; 99WO-US27147.
XX
XX 15-DEC-1998; 98US-0112332.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Mahajan PB;
XX WPI: 2000-452026/39.
XX P-PSDB: AAY95308.
XX
XX Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX
XX Example 1; Page 73-74; 85pp; English.
XX
XX The present sequence is that of maize cDNA (ATCC PTA-533) coding
CC for Rad2/FEN-1 (see AAY95308). The corresponding RNA was isolated
CC from a B73 line seedling after a 10 day drought, heat shocked for
CC 10 hr, and allowed to recover under normal conditions. Rad2/FEN-1
CC is a structure specific endonuclease which under certain conditions
CC also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used
CC to produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX Sequence 1541 BP; 473 A; 308 C; 377 G; 383 T; 0 other;
SQ
```

```
QY 19 CCGCGTTTCTTGGCCACTCCGGCTAGCCGCCGCCGCCGCCGCCGCCGCCGCCGCA 78
    |
Db 13 CGCGGTTTCTTGGCCACTCCGGCTAGCCGCCGCCGCCGCCGCCGCCGCCGCCGCA 72
QY 79 GACGAGATGGGCATCAAGGCTTTGACGAACCTGCTGGCGACATGGCCCAAGGGCATG 138
    |||||||
Db 73 GACGAGATGGGCATCAAGGCTTTGACGAACCTGCTGGCGACATGGCCCAAGGGCATG 132
QY 139 AAGGAGCAGAAAGTTCGAGAGCTTCTGGCCGCAAAATCCCGTCGACGCCAGCATGAC 198
    |||||||
Db 133 AAGGAGCAGAAAGTTCGAGAGCTTCTGGCCGCAAAATCCCGTCGACGCCAGCATGAC 192
QY 199 ATATACCAAGTTCCTGATTTGATTTGAGAGCAGAGCTGTAAGTCTGCAAAATGAGCT 258
    |||||||
Db 193 ATATACCAAGTTCCTGATTTGATTTGAGAGCAGAGCTGTAAGTCTGCAAAATGAGCT 252
QY 259 GGTGAGTCACTAGTCAATTTGCAAGGATGTTCAACCGGACATTAAGATTACTGAAAGCG 318
    |||||||
Db 253 GGTGAGTCACTAGTCAATTTGCAAGGATGTTCAACCGGACATTAAGATTACTGAAAGCG 312
QY 319 GGAATCAAGCCAGTTTATGTTTATGTTGATGGCAAGCCCTCTGATATGAAGAAACAGAGCTT 378
    |||||||
Db 313 GGAATCAAGCCAGTTTATGTTTATGTTGATGGCAAGCCCTCTGATATGAAGAAACAGAGCTT 372
QY 379 GCTAAAAGATFACCTCAAAAAGATGATGCAACCAAGATGCTGACTGAGGAGTACAGTA 438
    |||||||
Db 373 GCTAAAAGATFACCTCAAAAAGATGATGCAACCAAGATGCTGACTGAGGAGTACAGTA 432
QY 439 GGAGATTAAGATGCGATTAAGAAATTTGAGCAGAGAGCTGTAAGGTACAGAGGCAACAC 498
    |||||||
Db 433 GGAGATTAAGATGCGATTAAGAAATTTGAGCAGAGAGCTGTAAGGTACAGAGGCAACAC 492
QY 499 AAGCAAGATTTGAAAGGCTATTAAGCTATAGGGCTTCTGTTAAGAGACCTTCT 558
    |||||||
Db 493 AAGCAAGATTTGAAAGGCTATTAAGCTATAGGGCTTCTGTTAAGAGACCTTCT 552
QY 559 GAAGCAGAGCAGAAATGTCAGCCCTTTGCAATAACGATTAAGGTGTCCTGTTCCTTCA 618
    |||||||
Db 553 GAAGCAGAGCAGAAATGTCAGCCCTTTGCAATAACGATTAAGGTGTCCTGTTCCTTCA 612
QY 619 GAAGATATGAGACTCCCTTACTTTTGGGGCTCCACGGTCTCTGTCATTAATGATGCA 678
    |||||||
Db 613 GAAGATATGAGACTCCCTTACTTTTGGGGCTCCACGGTCTCTGTCATTAATGATGCA 672
QY 679 AGTCCCAAGAAATACCTGTGATGAAATTTGATGTTGCCAAGGTTTGGAGAGCTTGA 738
    |||||||
Db 673 AGTCCCAAGAAATACCTGTGATGAAATTTGATGTTGCCAAGGTTTGGAGAGCTTGA 732
QY 739 CTCACCATGACAGCATTCATGATTTGTCATCCTGTGATGATGATGATGATGATG 798
    |||||||
Db 733 CTCACCATGACAGCATTCATGATTTGTCATCCTGTGATGATGATGATGATGATG 792
QY 799 ATCAAAAGTATCGGGGGGCAACAGCTCTGAACCTTATTTGTCATGATGATGATG 858
    |||||||
Db 793 ATCAAAAGTATCGGGGGGCAACAGCTCTGAACCTTATTTGTCATGATGATGATG 852
QY 859 AAGCATCTTGAGAAATCTTAATAAAGACAGATATCAATTTCTGAGACAGGCTTACCA 918
    |||||||
Db 853 AAGCATCTTGAGAAATCTTAATAAAGACAGATATCAATTTCTGAGACAGGCTTACCA 912
QY 919 GAGCTCAGACCTGTTTTCAGGAGCTTAATGTCACATTTGATTTCTGAGCTAAATGG 978
    |||||||
Db 913 GAGCTCAGACCTGTTTTCAGGAGCTTAATGTCACATTTGATTTCTGAGCTAAATGG 972
QY 979 ACTGACCTGATGAGAGAGAGGCTCATAGTTTCTGTTAAACATATATGTTTCAACGAA 1038
    |||||||
Db 973 ACTGACCTGATGAGAGAGGCTCATAGTTTCTGTTAAACATATATGTTTCAACGAA 1032
QY 1039 GATCGGTGACAAAGGCCATGAGAGATCAATCTGCAAGATTAATATGTCGCAAGGA 1098
    |||||||
Db 1033 GATCGGTGAGAAAGGCCATGAGAGATCAATCTGCAAGATTAATATGTCGCAAGGA 1092
```


QY 889 TATCAATTCCTGAGAGCTGGCCCTTACCAAGAGCTGAGCGCTGTTCAAGAGCCTAAT 948
 |||||
 DB 841 TATCAATTCCTGAGAGCTGGCCCTTACCAAGAGCTGAGCGCTGTTCAAGAGCCTAAT 900
 QY 949 GTCACTTGGATATTCCTGAGCTAAATGACCTGACCTGATGAGAGGGTCTCATACT 1008
 |||||
 DB 901 GTCACTTGGATATTCCTGAGCTAAATGACCTGACCTGATGAGAGGGTCTCATAACT 960
 QY 1009 TTCCCTGTAAGAATATGTTTCAACGAGATCGGGTGCACAAGGCCATAGAGAATC 1068
 |||||
 DB 961 TTCCCTGTAAGAATATGTTTCAACGAGATCGGGTGCACAAGGCCATAGAGAATC 1020
 QY 1069 AAATCTGCCAAGATTAATGCTGCGAAGAGACTGAGTCTTTTTCAGCCCACTGCC 1128
 |||||
 DB 1021 AAATCTGCCAAGATTAATGCTGCGAAGAGACTGAGTCTTTTTCAGCCCACTGCC 1080
 QY 1129 ACCACATCAGACCGCTTAAACGAGAGAGACTTCGATTAACAAGCAGCAGCTGCG 1188
 |||||
 DB 1081 ACCACATCAGACCGCTTAAACGAGAGAGACTTCGATTAACAAGCAGCAGCTGCG 1140
 QY 1189 AACAGAAAACAAAGCTGCTGGAAGAGAAATTAATTTGGATGCTTATGTACACTA 1248
 |||||
 DB 1141 AACAGAAAACAAAGCTGCTGGAAGAGAAATTAATTTGGATGCTTATGTACACTA 1200
 QY 1249 CGACTAGAAAAGCAGCGGTGCGCTGATCACTTCTAGATTAATTAACCTGCTTTTA 1308
 |||||
 DB 1201 CGACTAGAAAAGCAGCGGTGCGCTGATCACTTCTAGATTAATTAACCTGCTTTTA 1260
 QY 1309 ACTAGAGCTTTGTTAAAGTTTGTCTCATGTTTCAAGCTGGGTAAGTTAGTTGTTT 1368
 |||||
 DB 1261 ACTAGAGCTTTGTTAAAGTTTGTCTCATGTTTCAAGCTGGGTAAGTTAGTTGTTT 1320
 QY 1369 AAGAGATTGTTGTACCAAGTACAAAACCTTATGCTGTTT 1411
 |||||
 DB 1321 AAGAGATTGTTGTACCAAGTACAAAACCTTATGCTGTTT 1363

RESULT 5
 AAX02117
 ID AAX02117 standard; DNA; 1144 BP.
 XX
 AC AAX02117;
 XX
 DT 23-APR-1999 (first entry)
 XX
 DE Human FEN-1 DNA fragment.
 XX
 KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasma; antineoplastic agent; cleavage; ss.
 XX
 OS Homo sapiens.
 XX
 FN US5874283-A.
 XX
 PD 23-FEB-1999.
 XX
 PD 30-MAY-1995; 95US-0455968.
 XX
 PF 30-MAY-1995; 95US-0455968.
 XX
 PR 30-MAY-1995; 95US-0455968.
 XX
 PA (HARR/) HARRINGTON J J.
 PA (HSIEH/) HSIEH C.
 PA (LIEBER/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 DR WPI; 1999-179985/15.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 PS Claim 3; Column 25; 58pp; English.
 XX

CC This sequence encodes a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing transgenic nonhuman animals expressing the
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.
 XX

SO Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other;

Query Match 20.7%; Score 302.6; DB 20; Length 1144;

Best Local Similarity 56.6%; Pred. No. 1.2e-66;

Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 85 ATGGGCAATCAAGGTTTGAAGAACTGCGGCGCAATGCGGCCCAAGCGCATGAAGAG 144
 |||||
 DB 1 ATGGGAATTCAGAGCCTGGCCAAACTAATTCGTATGCGCCCAAGTCCATCCGGAG 60
 QY 145 CAGAAGTTCGAGAGCTACTTGGCCGCAAAATCGCCGTGACGCGCAGCATATATAC 204
 |||||
 DB 61 AATGCATCAAGAGCTACTTGGCCGCAAAATCGCCGTGACGCGCAGCATATATAT 120
 QY 205 CAGTCTGATTTGATTTGGAAGAGACGATGGAACCTCACAATGAGCTGTGAA 264
 |||||
 DB 121 CAGTCTGATTTGATTTGGAAGAGACGATGGAACCTCACAATGAGCTGTGAA 177
 QY 265 GTCACTAGCTATTGCAAGAAATGTTCAACCGAGCAATAGATTACTGAAAGCGGAATC 324
 |||||
 DB 178 ACCACAGCCAGCTGATGAGCATGTTCTACCGCACATCGCATATGAGAGAGGCTATC 237
 QY 325 AAGCGATTATTTTGTATGAGCAAGCCCTGATGATGAAACAAACAGAGCTGTCTAA 384
 |||||
 DB 238 AAGCCCTGATTTCTTTGATGAGCAAGCCCGCAAGCTCAAGTCAAGGAGTGGCCAA 297
 QY 385 AGATCTCAAAAAGAGATGCAACCAAGATCTGACTGAGGAGTGAAGAGAT 444
 |||||
 DB 298 CGCAGTGAAGCGGCGGCTGAGCAGAGAGAGCTGACAGAGCTGAGCTGGGCGC 357
 QY 445 AAGATGCGGATTTGAAATTTGAGCAGAGGACGTGAAGCTGACAGGACACAGCA 504
 |||||
 DB 358 GAGCAGAGGTGAAATTTCACTAAGCGCTGTGAGAGGTACTAAGCAGACATGAT 417
 QY 505 GATTTAAAGCGCTATTAGACTTATGAGGAGTTCCTGTTGTAGAGCACTTCTGAACA 564
 |||||
 DB 418 GAGTCAAAACATCTGCTGAGGCTCATGAGGCTATCTTATCTTGAGCAACCAGTAGSCA 477
 QY 565 GAAGCAAAATGTCAGCCCTTGGCAATAAAGATAGAGTGTGCTGTGCTTCGCAAGAT 624
 |||||
 DB 478 GAGGCAAGCTGTGCTGCTGCTGCTGGAAGGCTGGCAAGGTATATGCTGCGCTACCGAGGAC 537
 QY 625 ATGACATCCCTTACTTTTGGGCTCCACGCTTCCTTCGATTTAATGATCAAGTTCC 684
 |||||
 DB 538 ATGACATCCCTTACTTTTGGGCTCCACGCTTCCTTCGATTTAATGATCAAGTTCC 597
 QY 685 AAGAAATTAACCTGTGATGAAATTTGATTTGTCAGAGTTTGGAGACCTGAGACTCAC 744
 |||||
 DB 598 AAAAGCTGCAATTCAGAGATTCACCTGAGCGGATTTCTCAGAGAGCTGGCTGAC 657
 QY 745 ATGACAGTTTCAATGATTTGTCATCTGTGTGATGATGATGATTTGATGATCAAAA 804
 |||||
 DB 658 CAGGACAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
 QY 805 GGTATCGGAGGCAACAGCTCTGAACATTTATTCGTAACATGGTCCATAGAAAGATC 864
 |||||
 DB 718 GGTATCGGAGGCAACAGCTCTGAACATTTATTCGTAACATGGTCCATAGAAAGATC 777
 QY 865 TTGAGATCTTAAATTAAGACAGATTCAAATTTCTGAGAGATGAGCTTACCAAGAGCT 924
 |||||
 DB 778 GTGCGGAGATTTGACCCCAACAAAGTACCTGTGTCAGAAATTTGGCTCCACAGAGAGCT 837

Db 1018 CTGCATGATTTCTTCAAGTGACCGGCTCACTCTCTTACAG 1058

RESULT 7
 AAX02111
 ID AAX02111 standard; DNA; 2033 BP.
 AC AAX02111;
 XX
 XX 23-APR-1999 (first entry)
 DE Human FEN-1 genomic DNA.
 XX
 XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
 KM neoplasma; antineoplastic agent; cleavage; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 104..1240
 FT /*tag= a
 FT /product= "FEN-1"
 XX
 XX US5874283-A.
 PN
 PD 23-FEB-1999.
 XX
 XX 30-MAY-1995; 95US-0455968.
 PF
 PR 30-MAY-1995; 95US-0455968.
 XX
 XX (HARR/) HARRINGTON J J.
 PA (HSIE/) HSIEH C.
 PA (LIEB/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX
 XX WPI; 1999-179985/15.
 DR P-PSDB; AAW92508.
 XX
 XX DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 PS Disclosure; Fig 5A-B; 58pp; English.
 XX
 XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing antibodies reactive with the novel
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.
 XX
 SQ Sequence 2033 BP; 500 A; 493 C; 579 G; 461 T; 0 other;

Query Match 18.9%; Score 276.6; DB 20; Length 2033;
 Best Local Similarity 55.9%; Pred. No. 66-60;
 Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 85 ATGGGCAATCAAGGTTTGAGAAACTGCTGGCGACATGCGCCCAAGCGGATGAAGAG 144
 DB 104 ATGGAATTCACGGCTTGCCAACTAATGCTGATGTGCCCCAGTGCATCCGTGAG 163
 QY 145 CAGAAGTTTCAGAGCTACTTCGCGCAAAATGCGCTGCAGCCAGCATGACATATAC 204
 DB 164 AATGACATCAAGAGCTACTTGTGCGAAAGTGCCATGCATGCTCCATGACATCTAC 223
 QY 205 CAGTTCCTGATGTAGTTGGAAGACAGCATGGAACCTCACAATGAGCTGGTGA 264
 DB 224 CAGTTCCTGATGTCTGTCAG--GmGGGATGTGCTGCAGAGACGAGAGGTGAG 280

QY 265 GTCACTAGTCATTTGCAAGCAATGTTCAACCGGACATTAAGATFACGAAAGCGGAATC 324
 DB 281 ACCACCAGCC---TGATGGGCAATGTTCTACCGTACCATGCG---CATGCAATATGGCATC 334
 QY 325 AAGCCAGTTTATGTTTATGATGCAAGCCTCCTGATATGAGAAACAGAGCTGTCTAAA 384
 DB 335 AAGCCTGTAGCTCTTTTATGCAACCAACAGCTAAGCTAAGAGAGCTGGCCAAAG 394
 QY 385 AGATCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCACTAGAGGTAGAGAT 444
 DB 395 CGCAGTGAAGAGCGCGCCGAGGCTGAGAAAGCAATGCAAGGCTCAGCAGCTGGATG 454
 QY 445 AAGATGCCATTTGAAAAATTTGAGCAGAGGAGCTGTAAGGTCACAAGCAACACAGNA 504
 DB 455 GAGGAGAGGTGGAAGATTCACCAAGAGGCTGTGTAAGTCAACCAAGACATCATGAT 514
 QY 505 GATTGTAACGGCTATTAAAGACTTATGGGGCTTCTGTTGTAGAGGACCTTCTGAAGCA 564
 DB 515 GAGTGCAAACACCTGCTGAGACCTCATGGGCTATCCCTTACCTTGATCACCAGCGAGCA 574
 QY 565 GAAGCAGATGTGACAGCCCTTTGCAATAACGATTAAGTGTGCGTGTGCTTCAGAGAT 624
 DB 575 GAGGCCAGCTGTGCTGCTGCCCTGGCAAGGCTGCAAAAGTATGCTCGCGCACGGAGGAC 634
 QY 625 ATGACATCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCATTTAAATGATCCAACTTC 684
 DB 635 ATGACATGCTCTACCTTTTGGCAGCCCGCTGCTAATGAGACATTAACCTGACATGAGGCC 694
 QY 685 AAGAAATACCTGTATGATGAATTTGATGTTGCCAAGTTTGGAGAGCTGTGAACCTACC 744
 DB 695 AAGAAGCTGCCATCCAAAGATTCATCTGAGCCGCTCTGAGAGAGGAGGCTGTGAAC 754
 QY 745 ATGACACATTCATGATTTTGGATCTGTCGTGATGATGATGATGATGATGATGATGAT 804
 DB 755 CAGGAGCAATTTGATGATCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
 QY 805 GGTATCGGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGGCTCATAGAAGCATC 864
 DB 815 GGCATTTGGCGCCAAAGCGGCTGTGATCTCATCCAGAAACATAGAAGCATGAGGATC 874
 QY 865 TTGAGAAATCTTAATAAAGACAGATATCAAAATTCCTGAGGAGCGGCTTACCAAGAGCT 924
 DB 875 GTGAGGGCGGCTGGACCCGACGACATCCCGTTCACAGAAATGCTGCCCAAGAGAGCC 934
 QY 925 GCAGCGCTTGTCAAGGAGGCTTAA--TGTCATATGATATTCCTGAGCTAAATGAGACT 981
 DB 935 CAGCAGCTCTTCTGAGACCGAAGTACTGACCCAGAGCTGTGAGGCTGGAAGTGAAGC 994
 QY 982 GCACCTGATGAGAGGCTCTCATAGTTTCTGCTGTAAGATTAATGCTTCAACGAAGAT 1041
 DB 995 GAGCCAAATGAAGAAGATGCTCAAAATTTATGTGTGTGTAAGAGCAAGTTTGTGAAGAG 1054
 QY 1042 CGGGTGACAAAGGCTATGAGAAATCAATTCGCCAAGAAATTAATGCTGCAAGAGAGA 1101
 DB 1055 CGAATTCGACGTGGGGGTAAACGGGTGAGTAAGAGCCGCCAGGCGACACCCAGAGAGCC 1114
 QY 1102 CTCGAGTCTTTTCAAGCAACTGCCACATCAGCACCCTGTAAGAGGAGAGACT 1161
 DB 1115 CTCGATGATTTTCTTAAGAGTGAAGGCTCAGCTCTCTAGC-----TAAAGCGC 1162
 QY 1162 TCGATTAACCAAGACAGGCTGCGAACAAGAAACAAAGGCTGGTGGAAACAGAAA 1221
 DB 1163 AAGGAGCCAGAAACCAAGGGGCTGCTAAAGAAAGAAAGACTGGGGGAGCGGGAGAG 1222
 QY 1222 T 1222
 DB 1223 T 1223

RESULT 8
 ABIL14287
 ID ABIL14287 standard; cDNA; 1300 BP.

```
XX ABL14287;
AC
XX
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37343.
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI: 2001-656860/75.
DR P-PSDB: ABB70184.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 37343; 21bp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 1300 BP; 346 A; 313 C; 369 G; 272 T; 0 other;
SQ
Query Match 18.9%; Score 276; DB 23; Length 1300;
Best Local Similarity 55.2%; Pred. NO. 7e-60;
Matches 581; Conservative 0; Mismatches 465; Indels 6; Caps 2;
QY 85 ATGGGCATCAAGGGTTGACGAACTGCTGGCGAGCATGCCCGCAAGCGCATGAAGAG 144
DB 1 ATGGGAATTTGGGCTTATCTAAGCTCATTTGCGCATCTGGCGCGACAGCGGATTCGGAA 60
QY 145 CAGAGTTGAGAGCTACTTGGCGCGCAAAATCGCGTTCAGCGCCAGCATGACATATAC 204
DB 61 AGTGAGATGAAGCAATTTTTCGTCGCAAGTACCAATCGATGTATAGTCCCTGTAC 120
QY 205 CAGTTCCTGATTGTAGTTGGAAGGACAGCATGGAATCTCACAATTAAGTGTGTA 264
DB 121 CAGTTCCTGATTCGCGCTCGCGTCCGAAGGC--GCCAGTTGGCGACCGTAATAGTGAT 177
QY 265 GTCACTATGCTATTCGAAGAATGTCAACCGGACATAAAGATTACTGGAAGCGGAAATC 324
DB 178 CCCAGCTCCCATTAATGCGCATGTCTACCGCACATCCGATTGCTGCAACAGGAAATC 237
QY 325 AAGCCAGTTTATGTTTGTATGAGCAAGCCCTCATATGAAGAACAAGAGCTTGCTAA 384
DB 238 AAGCCGATATATGCTTGTGAGCAAGCAACACACATCTCAAGTCCGCTGAGCTGGCGAAG 297
QY 385 AGATACTCAAAAAGAGATGATGACCAACCAAGATCTGACTGAGCGATGAGAGTACGAT 444
```

```
DB 298 CGGCCGAGCGCGGAGAGGAGCGAGAGGACACTGTAAGGCGCGCCACCGATCGGGAGAT 357
QY 445 AAGATGCGATTGAATAATGAGCAAGAGCATGTAAAGTCTCAAGCAACCAACGAA 504
DB 358 GATGCGGGTATCGAAAGTTTAATCGCCGATGGTCCGGGTAAGAGAGGACGCGCAAA 417
QY 505 GATTGTAACCGCTATTAGACTTATGGGGGTTCTGTTTGAAGGCACTTCTGAAGCA 564
DB 418 GAGGCGCAAGGAACTGCTACACCTAATGAGTGGTGCCCATGTTGATGCACCGCTCAAGG 477
QY 565 GAAGCAGATGTGCAAGCCCTTTGATTAAGCATTAAGGTGTCTCTGTTCTGACAAAT 624
DB 478 GAGGCCAGTGTGGCGCTGTGTGAAGGTAAGGTTTATGCTACCGCCAGCGAGGAT 537
QY 625 ATGACCTCCCTTACTTTTGGGGGCTCACGGCTTCCTGCTTGAATGAATCCAAAGTTC 684
DB 538 ATGGAATGCCCTACATTCGATGGAATCAAACTGTGAGTACTTACCTTAACAGCGAGCA 597
QY 685 AAGAAATATCTGTGATGGAATTTGATGTTGCCAAGGTTTGAAGAGCTTGAATCACC 744
DB 598 CGAAGATGCCCGTCAAGGAGATTCAGCTACGACAAAGCTGTGGAAGGTCTGGCCATTAA 657
QY 745 ATGACACAGTTCATTGATTTTGTGATCTCTGTGATGATGATGATGATGATGATGAT 804
DB 658 AATGACAGATTCATTGATCTATGATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 805 GGTATCGGGGGGCAACAGCTCTGAACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 864
DB 718 GGTATTTGACCCCAAGCGAGCGATGACATGATCAACCTATGCGGATATGAGACTATT 777
QY 865 TTGGAGATCTTAAATTAAGACATATCAATCTGTGAGGACTGGGCTTACCAAGAGCT 924
DB 778 CTGATTAACCTTGACTCTAGCAATATACCGCTGCCGAGAACTGGAACCTCAAGGTGG 837
QY 925 CGAGCTTTGTCAGAGAGCTTAATGTCAC--ATTGATATTCCTGAGCTCAAAATGACT 981
DB 838 CGGGAACCTCTTATCAACGAGCGAGGTAGCTGAGCCGACCTCCTAGATCTCAAAATG 897
QY 982 GCACCTGATGAGGAGGGGCTCATAGTTCTGCTGTAAGATTAAGATTAAGTCAAGCA 1041
DB 898 GAGCGGATGAGAGGAGGCGCTTGTCAAGTTCTGCGCGGACCGGACGTTCAACGAGAG 957
QY 1042 CGGTGCAAAAGCCATAGAGAGATCAAAATTCGCAAGATAAATGTCGCAAGAGA 1101
DB 958 CGCGTTCCGAACGCTGCGCAAAATCAAGCTGATGAATCCAGCAGGCCAGACTCAGTG 1017
QY 1102 CTCGAGTCTTTTCAAGGCCAAGTCCACAC 1133
DB 1018 CTCGATAGCTTTTAAAGACACTGCCACAGC 1049
RESULT 9
ID AAX02108 standard; cDNA; 1930 BP.
XX
XX AAX02108;
AC
XX 23-APR-1999 (first entry)
DE Mouse FEN-1 cDNA.
XX
XX FEN-1; mouse; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasia; antineoplastic agent; cleavage; ss.
XX
XX Mus sp.
OS
XX US5874283-A.
XX
XX 23-FEB-1999.
PD
XX 30-MAY-1995; 95US-0455968.
XX
```

```

PR 30-MAY-1995; 95US-0455968.
XX
XX (HARR/) HARRINGTON J J.
PA (HSIEH/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
XX Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI; 1999-179985/15.
DR P-PSDB; AAW92505.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
XX
PS Disclosure; Fig 2B; 58pp; English.
XX
XX This sequence encodes a mouse FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 1930 BP; 482 A; 459 C; 553 G; 436 T; 0 other:

Query Match      18.6%; Score 272.2; DB 20; Length 1930;
Best Local Similarity 55.1%; Pred. No. 7.6e-59;
Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;

QY 85 ATGGGCATCAAGGGTTTGAAGAACTGCTGGCGACATGCCCCCAAGGCATGAAGAG 144
   |||||  ||  ||  |||||  ||  ||  ||  |||||  ||  ||  ||
DB 1  ATGGGAATTCACGGCCTTCCAAACTAATGCTGATGTGGCCCCCAGTCCCATCGTAG 60

QY 145 CAGAAATTGAGAGCTACTTGGCGCCCAAAATGCGCGTGCAGCCAGTCGATATAC 204
   ||  ||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 61 AATGACATCAAGAGCTACTTGTGCTTAAAGTGCCATGATGCTCCATGAGCATCTAC 120

QY 205 CAGTTCCTGATTTGATTTGAAGAGCAGCATGAAACTCTCAAAATGAAAGCTGTGAA 264
   |||||  |||||  |||||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 121 CAGTTCCTGATTTGCTGTGCTGAG---GGTGGGATGCTGCTCAGAACGAGAGGTGAG 177

QY 265 GTCACTAGTCATTTGCAAGGAATGTTCACCGGACATTAAGTTACTGGAACGGGAATC 324
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 178 ACCACCAAGCCTGATGGCATGTTAT-----GCCAACCATCCGATGGAGAATGGCATC 231

QY 325 AAGCCAGTTTATTTTGTGATGGCAAGCCTCGATATGAAAGAAACAAAGCTTGCTAAA 384
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 232 AAGCTGTGTACTCTTTTGTATGCAAAACCAACCACTGAAGTCAGCGGAGGTGGCAAG 291

QY 385 AGATTACTCAAAAAGAGATGTCAGCAACCAAGATCTGATGAGGAGCTAGAGTAGAGAT 444
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 292 CGCAGTGAAGAGCGCGCGCGCTGAGAAAGCAACTGCACAGGCTCAGAGAGCTGGGATG 351

QY 445 AAGATGCTGATTAATAAATGAGCAAGAGAGCTGTAAGGTGCAGAGCAACAGCAAGAA 504
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 352 GAGGAGAGAGGTGAGAGGTTACCAAGAGGCTGTGAAGGTCCACAGCAACCAATGAT 411

QY 505 GATTGTAACGGCTATTATGACTTATAGGGGTTCTGTTGTAGAGGCACCTTCTGAAGA 564
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 412 GAGTGCAAAACCTCGTGAAGCTTCATGGGATCCCTTACTGATGACACAGCGAGGCA 471

QY 565 GAAGCAGAGATGTGAGCCCTTTTGATTAAGATAGAGTGTGCTGTGCTTCAAGAAAT 624
   ||  ||  |||||  |||||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 472 GAGGCCAGCTGTGCTGCTGGCGCAAGGCTGCAAAAGTATATCTGCGGACAGGAGAC 531

QY 625 ATGACATCCCTTACTTTTGGGGCTCCACGGTCTCTTGATTTAATGATCCAAATTC 684
   |||||  |||||  |||||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 532 ATGAGCTGCTCACTTTTGGCAGCCCGTGTCTAATGCGACTTAACCTGCAATGAGGCC 591

```

```

QY 685 AAGAAATACCTGTGATGGAATTTGATTTGGCCAAAGCTTTTGAGAGAGCTTGAACTCAC 744
   |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  ||  ||
DB 592 AAGAGCTGCGCCATCCAAAGTTTCATCTGAGCCGCGTCTCTCAGAGAGTGGTGTGAA 651

QY 745 ATGACACAGTTCAATTTGTCATCTGTGTGATGTGATGATTTGATGATCAAA 804
   |||||  |||||  ||  |||||  |||||  ||  |||||  |||||  |||||
DB 652 CAGGACAGATTTGTGATCTGTGCATCTCTGTGGTAGGACTACTGCGAGACATCCGT 711

QY 805 GGTATGGGGGGGAAACAGCTCTGAACCTTATTCGTAACATGGGCTCAAGAAAGCATC 864
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  ||  |||||
DB 712 GGCATTTGGCGCCAAAGCGGCTGTGATCTCATTCAGAAACATTAAGCATCAAGAGATC 771

QY 865 TTGAGACATCTTAATTAAGACAGATATCAAAATTCCTGAGAGCTGGCTTACCAAGAGCT 924
   ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  |||||  |||||
DB 772 GTGAGCGGCTGACCCCAAGTACCCCGTCCAGAAACATCGCTCCACAAGGAAGCC 831

QY 925 CGACGCTTTGCAAGAGCCTAA---TGTCACTTGGATTTCTGTAGCTAAATGAGACT 981
   ||  ||  |||||  |||||  ||  ||  ||  ||  |||||  |||||  |||||
DB 832 CAGCAGCTCTTCTGAGAGCGCAGAAAGTAGTGAGACCAGAGCTGTGTGAGATGAGAGC 891

QY 982 GCACCTGATGAGAGGGTCTCATAGTTTCTGTAAAGATATGTTTCAACGAAAT 1041
   ||  ||  |||||  |||||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 1042 CGGGTGCAAAAGGCCATAGAGATCAAAATCTGCCAAGATTAATCTGCCAAGAGAGA 1101
   ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 952 CGAATTCGAGTGGGGCTCAAGCGGCTGAGTAAGAGCCGCGCAGGAGCAGCCAGGGAGCC 1011

QY 1102 CTGCACTCTTTTTCAGGCACTGCCACATTCAGCACCCTTAAGGAAAGAGACT 1161
   |||||  |||||  |||||  ||  ||  ||  ||  ||  ||  |||||  |||||
DB 1012 CTCATATATTTCTCAAGGTGACAGGCTCACTCTCTCAGC-----TAAAGCCG 1059

QY 1162 TCGGATTAACAAAGCAGGACGCTGCGAACAGAAACAAAGGCTGTGTGAAAGAA 1221
   ||  ||  ||  ||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1060 AAGAGCGCAGAACCCCAAGGGGCTCTTAAGAAAGCAAAAGACTGGGAGAGGGGAAG 1119

QY 1222 T 1222
DB 1120 T 1120

RESULT 10
AA02109
ID AA02109 standard; cdna; 1149 BP.
XX
XX AA02109;
AC
XX 23-APR-1999 (first entry)
DT
XX
DE Yeast FEN-1 cdna.
XX
XX FEN-1; yeast; flap endonuclease; detection; diagnosis; carcinogen;
XX neoplasia; antineoplastic agent; cleavage; ss.
XX
XX Saccharomyces cerevisiae.
OS
XX
XX US5874283-A.
PN
XX
XX 23-FEB-1999.
PD
XX
XX 30-MAY-1995; 95US-0455968.
PE
XX
XX 30-MAY-1995; 95US-0455968.
PR
XX
XX (HARR/) HARRINGTON J J.
PA (HSIEH/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
XX Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI; 1999-179985/15.
DR
DR P-PSDB; AAW92506.
XX

```


PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 PS Disclosure; Fig 3B; 58pp; English.
 XX

CC This sequence encodes a yeast FEN-1 (flap endonuclease) protein. This
 CC protein is used in a method to isolate novel human FEN-1 proteins for
 CC detecting a pathological condition in a patient, for diagnostic purposes,
 CC for screening for antineoplastic agents and carcinogens, for diagnostic
 CC staging of neoplasia, for producing recombinant flap endonuclease for use
 CC as research or diagnostic reagents, for producing antibodies reactive
 CC with the novel polypeptides, for producing transgenic nonhuman animals
 CC expressing the novel polypeptides encoded by a transgene. The invention
 CC also provides a novel molecular cloning techniques and reagents involving
 CC cleavage of a flap or nick with a flap endonuclease.

SQ Sequence 1149 BP; 410 A; 182 C; 279 G; 278 T; 0 other;

Query Match 18.5%; Score 270.2; DB 20; Length 1149;
 Best Local Similarity 55.5%; Pred. No. 2e-58;
 Matches 599; Conservative 0; Mismatches 453; Indels 27; Gaps 3;

QY 85 ATGGGATCAGAGGGTTTGAGCAAACTGCGCGGACATCGCCCAAGCGATGAAGAG 144
 DB 1 ATGGGATTAAGAGTTTGATGCAATTAATATGGAACATGTTCCCTGCTACAGAAA 60

QY 145 CAGAAATTCGAGAGCTACTTGGCGCCCAAAATCGCCGTCGACGCAATGATCATATAC 204
 DB 61 AGGATATCAAGAGCTTTTGGCAGAAAGTTGCCATGATGATCCCTATGCTCATAT 120

QY 205 CAGTTCTCTATTTGATTTGGAAGACAGCATGGAACCTCTACAAATGAAGCTGTGAA 264
 DB 121 CAGTTTATTTATGCTGTAGACACAGCAAGCGGTGGCAGTTGACCAATGAACGCGTGAA 180

QY 265 GTTCACATGATTTGCAAGAAATGTTCAACCGGACATTAAGTTACGTGAGAGGGGAATC 324
 DB 181 ACAACGTCACTGATGGGTATGTTTATAGACACTGGAATGATGTAAGGTATC 240

QY 325 AAGCCAGTTTATGTTTGTATGGCAAGCCTCCGATATGAGAAACAGAGCTTGTAA 384
 DB 241 AAGCCTTTGATGTTGCTTCGACGCAAACTCCAGCTTTGAATCTGATGTCAGCAAG 300

QY 385 AGATACTCAAAAAAGATATGCAACCAAGATCTGACTGAGCACTAGAGTAGAGAT 444
 DB 301 CGCTCTTCAAGAGGTTGGAACAGAAAAAACTGCGACAGCAACACAGATTGAA 360

QY 445 AAAGATCGATGAAATTAAGACAGAGACTGTAAGGTGCAAGGCAACAGCAAGAA 504
 DB 361 AAGATGACAGCAAGAAAG-----AGATTGTTGAAGGTCTCAAAAGGCAATATGAA 411

QY 505 GATTGTAACGGCTATTAAGACTTATGGGGGTTCTGTTGTAGAGGCACCTTCTGAAGCA 564
 DB 412 GAAGCCCAAAATTAATAGCACTAATGGGAATCCCATATATATAGCGCAAGAGGCT 471

QY 565 GAAGCAGATGTGCAAGCCCTTTCATAAACGATAGGTGTGCTTGTCTTCAAGAT 624
 DB 472 GAGGCTCAATGTCGTAGTTGGCAAGAGGAAAGGTATGTCCTCCACAGTGAAGAT 531

QY 625 ATGAGACTCCCTACTTTTGGGCTCCAGCGTTCCCTCGCATTAATTAATGATCCATTC 684
 DB 532 ATGAGACACTCTGTTATAGAACACCCCTCTGTTAGACATTTGACTTTTGAAGGCC 591

QY 685 AAGAAATATCCTGTGATGAATTTGATGTGCAAGGTTTGGAGAGGCTTGAACATCACC 744
 DB 592 AAGAAAGAACCCGATTCACGAATATGACTGAATGTTTGGAGAGGCTCGACTTGACA 651

QY 745 ATGAGCAATTTATGATTTGTCATCTGTTGATGTCATTTGATGATGATCAAA 804
 DB 652 ATAGAGCAATTTGATGTTGTCATTAATGCTGTTGTCAGTACTGTAAAGCATCAGA 711

QY 805 GGTATCGGGGGCAACAGCTCTGAACATTTATTCGTCAACATGGGCTCAATGAACATC 864
 DB 712 GGTGTTGGTCAGTACAGCTTAATAATGATTAATAACGATGATTCATCGAAAAATC 771

QY 865 TTGAGATCTT-----ATTAAGACAGATATCAATTCCTGAGACTGG 909
 DB 772 GTGAGATTATTTGATCTGGGAGTCAAAACACATTAATGAAATCCAGAACTGG 831

QY 910 CCTTACCAAGACCTCGACGCTTTTCAGAGAGCCATTAATCTCATTTGATA--TTCT 966
 DB 832 CTTTACCAACCAAGAAATGCTTTCTTGACCCGTAAGTTATGATGTAAGCAATA 891

QY 967 GACCTAAATAGACTGACTGATGAGAGAGGCTCTCATAGTTTCCGTGTAAGATAAT 1026
 DB 892 AACTTGAAATGTCGCCACCAAGAGAAAGAACTTATCGAGATTTATGTGATGATAG 951

QY 1027 GATTTCACAGAAATCGGTCGACAAAGCCCATAGAGAAAGATCAATCTGCCAAGATAA 1086
 DB 952 AAATTCAGTGAAGAAAGATTAATCTGTTATATCAAGATTGAAAAAGCTTGAATCT 1011

QY 1087 TCGTCGCAAGGAAGACTCGAGTCTTTTTCAGAGCCAACTCCACACATCAGACCGCT 1145
 DB 1012 GGCATTCAAGGATAGTTAGATGGGTCTCTTCCAGTGTGCTTAAGACAAAGAAAGCT 1070

RESULT 11
 ID ABL14286
 AB14286 standard; cDNA; 3362 BP.
 XX
 AC ABL14286;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37340.
 XX
 KW Drosophila: developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 OS
 OS Drosophila melanogaster.
 PN
 PN WO200171042-A2.
 PD
 PD 27-SEP-2001.
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 PI
 DR MPI; 2001-656860/75.
 XX
 DR P-PSDB; ABB70183.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Claim 1. SEQ ID NO 37340; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pt_sequences.
 XX
 SQ Sequence 3362 BP; 979 A; 709 C; 750 G; 924 T; 0 other;

Query Match	17.3%;	Score 252.6;	DB 23;	Length 3362;
Best Local Similarity	55.2%;	Pred. No. 8.8e-54;		
Matches 536;	Conservative 0;	Mismatches 429;	Indels 6;	Gaps 2;
QY	166	GGCCGCAAAATCGCGTCGACGGCCAGCAGATATACCATGTTCTGATTTGAGTTGA	225	
DB	1144	GGTCGCAAGTACCAATCGATGCTAGTATGCTGTACCAATGCTCATCCGCGCCG	1203	
QY	226	AGACAGGATGGAACACTCACAATGAAAGCTGTGAACTCTACTACTCTTTGCAAGGA	285	
DB	1204	TCGGAAGGC---GCCAGTTGGCCACCGTAAATGATGTCACCGTCCCAATATGGGC	1260	
QY	286	ATGTTCAACCGCAGCAATTAAGTACTGGAAGCCGAATCAACGACAGTTATGTTTGTAT	345	
DB	1261	ATGTTCTACCGCAGCATCCGATTTGCTGAGCAAGGATCAAGCCGGATATGCTTCGAC	1320	
QY	346	GGCAGCCTCTGATATGAAACAAAGAGCTTGTAAAGATCTCAAAAGAGATGAT	405	
DB	1321	GGAAAGCCACAGATCTCAAGTCGGTGAAGCTGGCGAAGCGCCGAGCGCGAGGAA	1380	
QY	406	GCAACCAAGATCTGACTGAGCGAGTAAAGATGAGATTAAGTGTGAAATATG	465	
DB	1381	GCGGAGAGGACTGAAAGGCGCCACCGATGCGGAGATATGCGGTATCGAAACTTT	1440	
QY	466	AGCAAGAGACTGTAAAGTACAGAGCAACAGACAGATTTGTAAACGCTATTAGA	525	
DB	1441	AATCGCGATTTGGTCCGGTAAAGAGAGAGAGCCAAAGAGCGCAAGTGTCTACA	1500	
QY	526	CTTATGGGGTTCCTGTTTGAAGGACCTTCTGAGCAGACAGATGTCAGCCCTT	585	
DB	1501	CTATATGGTGTGCGCTATGTTGATGACCCGTGGAGCGGAGCGCCAGTGTGCGCTGT	1560	
QY	586	TGATTAAGCATTAAGTGTGCTGCTTCAAGATATGACCTCCCTTACTTTGGG	645	
DB	1561	GTCGAAGCTGGAAGGTTTATGCCACCGCCAGAGAGATTTGGATGCCCTTCACATTGGA	1620	
QY	646	GCTCCAGGTTCTCTGCTCATTTAATGATCAAGTTCAGAAATATCTGTATGAA	705	
DB	1621	TTTACAAACTGTTGAGATACCTTACTACAGGAGCAGAAAGATGCCGTCACAGAG	1680	
QY	706	TTTGAATGTTCCCAAGTTTGGAGAGCTTGAACCTCACATGAGCCAGTTCATTGATTG	765	
DB	1681	TTAGCTACACAAAGTGTGGAAGGCTGCGCATTAACAATCGAGAGTTCATTGATCTA	1740	
QY	766	TGATCTCTGTGTGATGTGACTATTTGATAGCATCAAGGTATCGGGGGCAAAAGCT	825	
DB	1741	TGATTTCTGCTGGTTGCCATTTACTGTGAGAGCATCAAGGATTTGACCCCAAGCGG	1800	
QY	826	CTGAACCTTATTCGTCAACATGGGTCCATAGAAAGCATTTGGAGAATCTTAATAAGAC	885	
DB	1801	ATCGAAGTGTATCAACACTTATCGGATATAGAGACTATTCGTGATTAACCTGACTAGC	1860	
QY	886	AGATATCAATTCCTGTGAGCTGCGCTTACCAAGAAAGCTGAGCCCTGTTCAGAGACCT	945	
DB	1861	AAATCAACCGTGGCCGGAAGTGAAGTGAAGTGGCGGGGAAGCTCTTCATCGAACCG	1920	
QY	946	AATGTAC---ATTGATATTCCTGAGCTAAATGAGCTGAGCTGATGAGAGGCTCTC	1002	
DB	1921	GAGGTACTGATGCGGCTCCATATGATCTCAATGGGTGCGAGCGGATGAGAGGCGCTT	1980	
QY	1003	ATTAAGTTCTCTGGTAAAGATATGTTTCAAGAGATGGGTGACAAAGCCATAGAG	1062	
DB	1981	GTCGAAGTTTCTGCGCGCGAGCGGCAATCAACGAAAGAGGCGCTTCCACAGCGTGCCAAA	2040	
QY	1063	AAGATCAATCTGCGCAGATTAATGTCGCAAGAGAGACTCGAGTCTTTTCAAGCA	1122	
DB	2041	AAGGTGTAATCAAGCAGGCGCCAGACTAGGTGAGACTCGTACTCTTTAAGACA	2100	
QY	1123	ACTGCCACAC 1133		
DB	2101	CTGCCACGAC 2111		

RESULT 12	
ABL10094	
ID ABL10094 standard; cDNA: 5471 BP.	
XX	
AC ABL10094;	
XX	
DT 26-MAR-2002 (first entry)	
XX	
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 24764.	
XX	
KW Drosophila; developmental biology; cell signalling; insecticide;	
KW pharmaceutical; gene; ss.	
XX	
OS Drosophila melanogaster.	
XX	
PN WO200171042-A2.	
XX	
PD 27-SEP-2001.	
XX	
PF 23-MAR-2001; 2001WO-US09231.	
XX	
PR 23-MAR-2000; 2000US-191637P.	
PR 11-JUL-2000; 2000US-0614150.	
XX	
PA (PEKE) PE CORP NY.	
PI Venter JC, Adams M, Li PWD, Myers EW;	
PI WPI; 2001-656860/75.	
DR P-PSDB; ABB65991.	
XX	
PT New isolated nucleic acid detection reagent for detecting 1000 or more	
PT genes from Drosophila and for elucidating cell signalling and cell-cell	
PT interactions -	
XX	
PS Claim 1: SEQ ID NO 24764; 21pp + Sequence Listing; English.	
XX	
CC The invention relates to an isolated nucleic acid detection reagent	
CC capable of detecting 1000 or more genes from Drosophila. The invention is	
CC useful in developmental biology and in elucidating cell signalling and	
CC cell-cell interactions in higher eukaryotes for the development of	
CC insecticides, therapeutics and pharmaceutical drugs. The invention	
CC discloses genomic DNA sequences (ABL1840-ABL16175), expressed DNA	
CC sequences (ABL1840-ABL16175) and the encoded proteins	
CC (ABB57737-ABB72072).	
CC The sequence data for this patent did not form part of the printed	
CC specification, but was obtained in electronic format directly from WIPO	
CC at ftp.wipo.int/pub/published_pcl_sequences.	
XX	
SQ Sequence 5471 BP; 1422 A; 1223 C; 1325 G; 1501 T; 0 other;	
XX	
Query Match	11.8%; Score 173.2; DB 23; Length 5471;
Best Local Similarity	54.0%; Pred. No. 1.4e-33;
Matches 377; Conservative	0; Mismatches 318; Indels 3; Gaps 1;
QY	439 GGAGATTAAGATGGATTGAATAATGAGCAAGAGACTGTAAAGTACACAGGCAACAC 498
DB	4 GGAATGATGCCGGTATCGAAAGTTTAATCGCGGATTTCCGGGTAAAGAGGAGACAC 63
QY	499 AACGAATTTGTAACCGCTATTAGACTTATGGGGTCTCTGTTGAGAGGACCTTCT 558
DB	64 GCCAAAGAGGCCAAGAGACTGCTACACTAATGGGTGCTGCTATGTTGATGACGCTGC 123
QY	559 GAAGCAGAAAGCAAGATGTCAGCCCTTTGATTAACGATTAAGGATGTTGCTGTTCTTCA 618
DB	124 GAACGAGAGGCCAGTGTGCCGCTCTGTGTAAGCTGGAAGGTTTATGCGACGCCACG 183
QY	619 GAAGATATGAGACTCCCTTACTTTTGGGCTCCACGCTTCTCTGTCATTTAATGATCA 678
DB	184 GAGGATATGAGATGCCCTCACATTCGATTCGTAACAAACTGTTAGATACCTTACTACGC 243
QY	679 AGTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCCAAGCTTTTGGAGAGCTTGA 738

Db	244	GAGGACACGAAGATGCCCGTCAAGAGATTTCAGTACACAAAGCTGTTGGAAAGTCTGGCC	303
Qy	739	CTCACCATTGGACACGATTCATTGATTTGTGCATCCCTGTGTGATGTGATTTGTATAGC	798
Db	304	ATTAAACAATCGAAGATTCTATGATCTATGTATTCTGTGGGTGCATTACTGTAGAAC	363
Qy	799	ATCAAGAGTATCGGGGGGACAAACAGCTCTGAAACTTATTGTTCATAGATGGTCCATAGAA	858
Db	364	ATCAAGGGTATTGGACCAACCAACGACGATCGAACTGATCAACACTATCGGGATATAGAG	423
Qy	859	AGCATCTTGGAGAACTCTTAATAAAGACAGATATCAAAATTCCTGAGGACTGGCTTACCAA	918
Db	424	ACTATTCTGGATTAACCTGGACTTAGCAATTAACCGTGGCCGAGAACTGAAGATTAACAAG	483
Qy	919	GAAGTCGACCCCTTTTCAAGGAGCCTAATAGTCATCAATTTGATATTTCCT--GAGCTAAAA	975
Db	484	GTGGCGCGGGAACCTCTTATCGAAGCCGAGAGTGTGATGTGCGACTTCATTAATCTCAAA	543
Qy	976	TGGACTGCACCTGATGAGGAAGGTCCTCATAGTTTCCTGGTAAAGATATAGTTTCAAC	1035
Db	544	TGGGTGCGAGCCGATGAGGAGGCGCTTGTCAAGTTTCTCTGGGGGACCGGAGTTCAAC	603
Qy	1036	GAAGTCGGGTGCACAAAGGCCATAGAGATCAATTTGCCAAGATTAATCTGTGCGAA	1095
Db	604	GAGAGCGCGTTGCCCAACGGTCCCAAAAAGCTGATGAAATCCAAAGCAGGCCCACTAGC	663
Qy	1096	GGAGACTCGAGTCCCTTTTTCACGCCCAACTGCCACCAAC	1133
Db	664	GTGAGACTCGATAGTACTTCTTAAGACACTGCCCCAGAC	701
RESULT 13			
AAAT76685			
ID	AAAT76685	standard; DNA; 1023 BP.	
XX	AAAT76685;		
DT	14-APR-1998	(first entry)	
DE	Pyrococcus furiosus FEN-1 endonuclease gene coding sequence.		
XX			
KW	Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;		
KW	invader directed cleavage; FEN-1; endonuclease; ds.		
OS	Pyrococcus furiosus.		
PN	W09727214-A1.		
XX	31-JUL-1997.		
XX			
PF	22-JAN-1997;	97MO-US01072.	
XX			
PR	02-DEC-1996;	96US-0759038.	
PR	24-JAN-1996;	96US-0599491.	
PR	12-JUL-1996;	96US-0682853.	
PR	29-NOV-1996;	96US-0756386.	
PR	02-DEC-1996;	96US-0758314.	
XX			
PA	(THIR-) THIRD WAVE TECHNOLOGIES INC.		
XX			
P1	Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI;		
P1	Olive DM, Prudent JR;		
XX			
XX			
DR	WPI: 1997-393613/36.		
DR	P-PSDB; AAM24216.		
XX			
PT	Thermostable structure-specific nuclease(s) - used for detection and		
PT	characterisation of nucleic acid sequences and variations in nucleic		
XX	acid sequences		
XX			
PS	Example 28: Page 283-285; 457pp; English.		
XX			
CC	This sequence comprises the coding region of the gene encoding		

CC Pyrococcus furiosus (Pfu) FEN-1 endonuclease (see A0424216).
CC It was obtained by PCR amplification (see AAT76682-83). Large
CC scale preparation of recombinant Pfu FEN-1 was performed using
CC *E. coli* as host. Pfu FEN-1 is a thermostable enzyme. It can be
CC used in novel methods for the detection and characterisation of
CC nucleic acid sequences and variations in nucleic acid sequences.
XX
50 Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other.

Query Match	10.9%;	Score 159.6;	DB 18;	Length 1023;
Best Local Similarity	54.6%;	Pred. No. 1.9e-30;		
Matches 318;	Conservative 0;	Mismatches 264;	Indels 0;	Gaps 0;

QY	134	CGATGAGGACACAGATCTGGAGAGCTACTTCGGGCCCAAAATCCGCTGCAGCCAGCA	193
Db	29	CAAGAAAAGAAATTGAATTGAAAAACCTTAATCGGGAATAAAATTCGCAATGACGCTCTTA	88
QY	194	TGACCATATACCACTTCCTGATTGCTTACCTTGGAAAGACAGGCATGGAACCTCCAAATG	253
Db	89	ATGCATCTACCAATTTTGTGCCAATATAGACGAATAAGATGCAATCCACTTATGATT	148
QY	254	AAGCTGGTGAAGTACTATGTCATTTTGCAGAGAAATGTTCAACCGACATATAGATTACTGG	313
Db	149	CAAAGGGTAGAATTAACCTCCACACTAAGCGGGCTTTTACAGACAAATTAACCTAATGG	208
QY	314	AAGCGGAAATCAACGACGTTATCTTTTATGTCGCAAGCCTCCTGATATACAAACACAG	373
Db	209	AGGCTGGAAATTAACCTCTGTATGTTTTTATGAGAACTCCAGATTCATAAAAGAAAC	268
QY	374	AGCTTGTAAAGATTACTCAAAAAGAGATATGCAAACAAAGATCTGACTAGACGAGTAG	433
Db	269	AGCTCGAAAAAAGAAAGAGAGCGGACGAGAGACCTGAAAGAAAGTGGAGGAAGCACTTG	328
QY	434	AGGTAGAGATTAAGATGCGATTGAAAAATTGACAGCAAGGACTGTAAAGTCACAAGGC	493
Db	329	AAAAAGAGATATAGAGAGAAAGAAATAATCCCAAAAGACAAACAGGTTAAATGAAA	388
QY	494	AACACAAGAGAAATGTAAACGCGTATTATAGCTTTATGGGGGTTCCGTGTATAGAGCAAC	553
Db	389	TGCTCATATCGAGGATGCAAAAAACTCTTAGAGCTTTATGGAAATCTCTATATGTTCAAGCAC	448
QY	554	CTTCTGGAAGCAGAAAGCAATGTGCAGCCCTTTGCATTAACGATTAAGAGTTCGCTGTGG	613
Db	449	CTAGCGGAGGAGAGGCCCAAGCTGCATATATGTGCCGGAATAAGGAGACGCTATAGCATGG	508
QY	614	CTTCAGAAATATGAGACTCCCTTACTTTTGGGCTCCACGGTTCCTTGCATTTAATGG	673
Db	509	CTAGTCAGATTTACGATTCCTCTACTTTTGGAGACTCCAAAGACTTGTAGAAACTTAACAA	568
QY	674	ATCCAAATCTCAGAAAAATCTCTGTGATGGAAATTTGATGTG	715
Db	569	TAACAGAAAAAACAAGTCTCTGGGAAAAAATGTTCTACGCTG	610

	RESULT 14
AAV65840	
ID	AAV65840 standard; DNA; 1023 BP.
XX	
AC	AAV65840;
XX	
DT	02-FEB-1999 (first entry)
XX	
DE	Pyrococcus furiosus FEN-1 endonuclease gene ORF.
XX	
KW	Nucleic acid detection; multiple sequential invasive cleavage;
KW	FEN-1; endonuclease; nuclease; ds.
XX	
OS	Pyrococcus furiosus.
XX	
PN	WO9842873-A1.
XX	
PD	01-OCT-1998.
XX	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:33:28 : Search time 35 8811 Seconds
(without alignments)
10015.359 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463

Sequence: 1 caccgaatagctcgccgcgc.....aaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 segs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backlist.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1463	100.0	1463	4	US-09-426-557-1
2	1421.8	97.2	1478	4	US-09-426-557-7
3	1409.8	96.4	1541	4	US-09-426-557-3
4	1350.2	92.3	1381	4	US-09-426-557-5
5	302.6	20.7	1144	2	US-08-455-968E-2
6	302.6	20.7	1144	2	US-08-455-968E-28
7	276.6	18.9	2033	2	US-08-455-968E-9
8	272.2	18.6	1930	2	US-08-455-968E-4
9	270.2	18.5	1149	2	US-08-455-968E-6
10	159.6	10.9	1023	2	US-08-757-653-175
11	159.6	10.9	1023	2	US-08-823-516-78
12	159.6	10.9	1023	3	US-08-759-038-114
13	159.6	10.9	1023	3	US-08-758-314-114
14	135.6	9.3	1032	4	US-09-146-319-1
15	135.6	9.3	1032	4	US-09-175-973-1
16	89.2	6.1	981	2	US-08-757-653-171
17	89.2	6.1	981	2	US-08-823-516-74
18	89.2	6.1	981	3	US-08-759-038-110
19	89.2	6.1	981	3	US-08-758-314-110
20	65.2	4.5	7218	1	US-08-232-463-14
21	47	3.2	455	1	US-08-636-928-4
22	47	3.2	1736	3	US-09-182-816-22
23	47	3.2	1736	3	US-09-182-816-22
24	47	3.2	1736	3	US-09-471-528-24
25	47	3.2	1736	3	US-09-471-528-24
26	47	3.2	1736	3	US-09-634-530-22
27	47	3.2	1736	4	US-09-634-530-24

C	28	46.4	3.2	685	4	US-09-227-357-66	Sequence 66, Appl
	29	46		3246	3	US-09-005-180A-2	Sequence 2, Appl
	30	45.4	3.1	1359	4	US-09-387-574-11	Sequence 11, Appl
	31	45.4	3.1	1359	4	US-09-668-096-11	Sequence 11, Appl
	32	45.2	3.1	873	4	US-09-475-316A-20	Sequence 20, Appl
	33	45.2	3.1	2205	3	US-08-888-077A-41	Sequence 41, Appl
	34	45	3.1	222	4	US-08-481-190-15	Sequence 15, Appl
	35	45	3.1	222	5	PCT-US93-00869-15	Sequence 15, Appl
	36	44.4	3.0	3410	4	US-09-020-956-110	Sequence 110, App
	37	44.4	3.0	3410	4	US-09-030-607-110	Sequence 110, App
	38	44.4	3.0	3410	4	US-09-439-313-110	Sequence 110, App
	39	43.8	3.0	1223	3	US-09-154-874-4	Sequence 4, Appl
	40	43.8	3.0	1534	4	US-08-300-903A-6	Sequence 6, Appl
	41	43.8	3.0	1882	4	US-09-370-253-1	Sequence 1, Appl
	42	43.6	3.0	893	1	US-08-276-452A-66	Sequence 66, Appl
	43	43.6	3.0	893	2	US-08-798-744-66	Sequence 66, Appl
	44	43.2	3.0	700	1	US-07-846-992-1	Sequence 1, Appl
	45	43.2	3.0	700	1	US-08-469-553-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-426-557-1
Sequence 1, Application US/09426557
Patent No. 6232527
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961
CURRENT APPLICATION NUMBER: US/09/426,557
CURRENT FILING DATE: 1998-10-22
EARLIER APPLICATION NUMBER: 60/112,332
EARLIER FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1463
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (85)...(1221)
US-09-426-557-1

Query Match 100.0%; Score 1463; DB 4; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CACGAGATAGCTCGCGCCGCGTTCTTTCGCGCTCCGCGTCAAGCGTTCGCGGAC	60
DB	1	CACGAGATAGCTCGCGCCGCGTTCTTTCGCGCTCCGCGTCAAGCGTTCGCGGAC	60
QY	61	CCCGCCACAGCCCGCGCGAGACATGGCATCAAGGTTTGACGAAACTGTGCGGAC	120
DB	61	CCCGCCACAGCCCGCGCGAGACATGGCATCAAGGTTTGACGAAACTGTGCGGAC	120
QY	121	ATTCGCCCAAGCGCATGAGAGAGAGAGTTCGAGAGCTACTTCGCGCCCAAAATCGCC	180
DB	121	ATTCGCCCAAGCGCATGAGAGAGAGAGTTCGAGAGCTACTTCGCGCCCAAAATCGCC	180
QY	181	GTTCGAGCCAGCATGATATACAGTCTCTGATTTAGTTGAAAGACAGCATGAA	240
DB	181	GTTCGAGCCAGCATGATATACAGTCTCTGATTTAGTTGAAAGACAGCATGAA	240
QY	241	ACTCTACAATGAAGCTGTGAAGTCACTAGTATTGCAAGGAATGTTCAACCGACA	300
DB	241	ACTCTACAATGAAGCTGTGAAGTCACTAGTATTGCAAGGAATGTTCAACCGACA	300
QY	301	ATTAAGTCTCTGAAGCGGAAATCAAGCACTTATGTTTGTGCAAGCCTCTGAT	360
DB	301	ATTAAGTCTCTGAAGCGGAAATCAAGCACTTATGTTTGTGCAAGCCTCTGAT	360

Db 301 ATAAAGTTACTGGAACGGGAATCAAGCCAGTTATGTTTTGATGGCAACCTCCTGAT 360
QY 361 ATGAAGAACACAGAGCTTCTAAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTG 420
Db 361 ATGAAGAACACAGAGCTTCTAAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTG 420
QY 421 ACTGAGGAGTGAAGTGAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 480
Db 421 ACTGAGGAGTGAAGTGAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 480
QY 481 AAGGTCAACAAGCAACACAGATGTAACGCGCTATTAGACTTATGGGGTTCCT 540
Db 481 AAGGTCAACAAGCAACACAGATGTAACGCGCTATTAGACTTATGGGGTTCCT 540
QY 541 GTTGATGAGGACCTTCTGAAGACAGAGATGATGAGGACCTTCTGAATGAAGATGAAG 600
Db 541 GTTGATGAGGACCTTCTGAAGACAGAGATGATGAGGACCTTCTGAATGAAGATGAAG 600
QY 601 GTGTCGCTGCTTCTGAGAGATATGATCCTTACTTTTGGGGCTCCACGGTTCCT 660
Db 601 GTGTCGCTGCTTCTGAGAGATATGATCCTTACTTTTGGGGCTCCACGGTTCCT 660
QY 661 CGTCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 CGTCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 721 GTTTTGGAGAGCTTGAACCTCAACATGACATGACATGATGATGATGATGATGATGATG 780
Db 721 GTTTTGGAGAGCTTGAACCTCAACATGACATGACATGATGATGATGATGATGATGATG 780
QY 781 TGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 781 TGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 841 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 901 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 901 GAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 ATTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 961 ATTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 AATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Db 1081 AATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 CCGGTAAAG 1200
Db 1141 CCGGTAAAG 1200
QY 1201 AAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 AAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 CAGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1261 CAGGAGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
QY 1321 GGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Db 1321 GGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TACCAAGTAAACAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1381 TACCAAGTAAACAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440

QY 1441 AAAAAAAAAAAAAAAAAAAAAA 1463
Db 1441 AAAAAAAAAAAAAAAAAAAAAA 1463

RESULT 2
US-09-426-557-7
Sequence 7, Application US/09426557
Patent No. 623527
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Maize Rad2/ERN-1 Orthologues and Uses
TITLE OF INVENTION: thereof
FILE REFERENCE: 0961
CURRENT APPLICATION NUMBER: US/09/426,557
CURRENT FILING DATE: 1999-10-22
EARLIER APPLICATION NUMBER: 60/112,332
EARLIER FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 1478
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (97)...(1233)
US-09-426-557-7

Query Match 97.2%; Score 1421.8; DB 4; Length 1478;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 7 AATAGCTGCGGCGCGGCTTCTGCGCCATCGCGGCTACGCGCGCCGCCACCCGCC 66
Db 19 AATAGCTGCGGCGCGGCTTCTGCGCCATCGCGGCTACGCGCGCCGCCACCCGCC 78
QY 67 ACAGCCGCCCGCAACAGAGATGGGATCAAGGCTTGGAGAAATGCTGGGAGCAATGCG 126
Db 79 ACAGCCGCCCGCAACAGAGATGGGATCAAGGCTTGGAGAAATGCTGGGAGCAATGCG 138
QY 127 CCCAAGGCGATGAAGAGACAGAAATGCGAGACTTCTGCGCGCAAAATGCGCGTGGAC 186
Db 139 CCCAAGGCGATGAAGAGAGCAAGAAATGCGAGACTTCTGCGCGCAAAATGCGCGTGGAC 198
QY 187 GCCAGCATGAGCATATACAGAGTCTGATGATGATGATGATGATGATGATGATGATGATG 246
Db 199 GCCAGCATGAGCATATACAGAGTCTGATGATGATGATGATGATGATGATGATGATGATG 258
QY 247 ACAATATGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 306
Db 259 ACAATATGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 318
QY 307 TTACTGGAACCGGGAATCAAGCCAGTTTATGTTTGAATGGCAAGCCCTCGATATGAAG 366
Db 319 TTACTGGAACCGGGAATCAAGCCAGTTTATGTTTGAATGGCAAGCCCTCGATATGAAG 378
QY 367 AAACAAGAGCTTGTCTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 426
Db 379 AAACAAGAGCTTGTCTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 438
QY 427 GCATAGAGTGAAGATGAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 486
Db 439 GCATAGAGTGAAGATGAAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 498
QY 487 ACAAGGCAACAACAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 546
Db 499 ACAAGGCAACAACAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 558
QY 547 GAGGACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 606
Db 559 GAGGACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 618

QY	607	GCCTGGCTCACAAGATATGAGCTCCCTTACTTTTGGGGCCACGGTCCCTGTCAT	666
Db	619	GCCTGGCTCACAAGATATGAGCTCCCTTACTTTTGGGGCCACGGTCCCTGTCAT	678
QY	667	TTAATGATCCAAAGTTCACAGAAATTAACCTGTGATGAATTTGATGTTGCCAAGGTTTGG	726
Db	679	TTAATGATCCAAAGTTCACAGAAATTAACCTGTGATGAATTTGATGTTGCCAAGGTTTGG	738
QY	727	GAGAGCTTGAACATCCACATGAGCCAGTTCATTGATTTGGATCCCTGTGATGTGAC	786
Db	739	GAGAGCTTGAACATCCACATGAGCCAGTTCATTGATTTGGATCCCTGTGATGTGAC	798
QY	787	TATTGTGATAGCTCAAGGTATCGGGGGGCAACAGCTCTGAACCTTATTCGTCAACAT	846
Db	799	TATTGTGATAGCTCAAGGTATCGGGGGGCAACAGCTCTGAACCTTATTCGTCAACAT	858
QY	847	GGGTCCATAGAAAGCATCTCTGGGATCTTAAATAAGACAGATATCAAAATTCCTGAGAC	906
Db	859	GGGTCCATAGAAAGCATCTCTGGGATCTTAAATAAGACAGATATCAAAATTCCTGAGAC	918
QY	907	TGGCGTTACCAAGAACCTGAGCGCTGTGTCAAGAGAGCCATAAGTCACTTGGATATTCCT	966
Db	919	TGGCGTTACCAAGAACCTGAGCGCTGTGTCAAGAGAGCCATAAGTCACTTGGATATTCCT	978
QY	967	GAGCTAAATAGACTCGACCTGATGAGAGGGCTCATTAAGTTTCCCTGTAAAGATAT	1026
Db	979	GAGCTAAATAGACTCGACCTGATGAGAGGGCTCATTAAGTTTCCCTGTAAAGATAT	1038
QY	1027	GGTTTCACGAABATGGGGGACAAAGGCCATAGAAAGATCAAAATCTGCCAAGATATA	1086
Db	1039	GGTTTCACGAABATGGGGGACAAAGGCCATAGAAAGATCAAAATCTGCCAAGATATA	1098
QY	1087	TCGTCCGAAGGAAGACTCGAGTCCCTTTTCAAGGCCAACATGCCACCACATCAGACCCGTA	1146
Db	1099	TCGTCCGAAGGAAGACTCGAGTCCCTTTTCAAGGCCAACATGCCACCACATCAGACCCGTA	1158
QY	1147	AAAGCGAAGAGAGCTTCGGATTAACAACCAAGGCAAGCTGGCAACAACAAAACAAAGGCT	1206
Db	1159	AAAGCGAAGAGAGCTTCGGATTAACAACCAAGGCAAGCTGGCAACAACAAAACAAAGGCT	1218
QY	1207	GGTGGAAAGAGAATATATCTTGATGCTTGAATGTACAACTACGACTACGAAAGCAGCGG	1266
Db	1219	GGTGGAAAGAGAATATATCTTGATGCTTGAATGTACAACTACGACTACGAAAGCAGCGG	1278
QY	1267	TGGCGTATACACTCCCTAGATTTATTTAACTCCCGTGTTAACCTCAAGAGCTTTGGTAA	1326
Db	1279	TGGCGTATACACTCCCTAGATTTATTTAACTCCCGTGTTAACCTCAAGAGCTTTGGTAA	1338
QY	1327	AGTTTGCTCATGTTTCAAGCTGGGGGTAAGTGTGTTTAAGAGATTGGTGTACCA	1386
Db	1339	AGTTTGCTCATGTTTCAAGCTGGGGGTAAGTGTGTTTAAGAGATTGGTGTACCA	1398
QY	1387	GTACAAACATATATCGCTGTTTTTAACTCTTGCTCTTGAAGATATAAAAAAATAAAAA	1446
Db	1399	GTACAAACATATATCGCTGTTTTTAACTCTTGCTCTTGAAGATATATATGACAGATAAA	1458
QY	1447	AAAAAAAAAAAAAAAAAAAA 1463	
Db	1459	AAAAAAAAAAAAAAAAAAAA 1475	

```

RESULT 3
US-09-426-557-3
: Sequence 3, Application US/09426557
: Patent No. 6232527
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
: TITLE OF INVENTION: Theoret
: FILE REFERENCE: 0961
: CURRENT APPLICATION NUMBER: US/09/426,557
: CURRENT FILING DATE: 1999-10-22

```

```
?  
? EARLIER APPLICATION NUMBER: 60/112,332  
? FILER FILING DATE: 1998-12-15  
? NUMBER OF SEO ID NOS: 10  
? SOFTWARE: FASTSQ for Windows Version 3.0  
? SEO ID NO 3  
? LENGTH: 1541  
? TYPE: DNA  
? ORGANISM: zea mays  
? FEATURE:  
? NAME/KEY: CDS  
? LOCATION: (79)...(1215)  
IS-09-426-357-3
```

Query Match	96.4%;	Score 1409.8;	DB 4;	Length 1541;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1411; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	19	CCGGGATTTTTCGCGCACGCGGCTCAGCCGCGCCGCGCACCGCACAGCGCGCGCA	78
Db	13	CGGGGATTTCTTCGCGCACGCGGCTCAGCCGCGCCGCGCACCGCGCACAGCGCGCGCA	72
QY	79	GAGCAGATGGGCACTCAAGGGTTTGAAGAACTGCTGGGGGACATTCGGGCCCAAGGCATG	138
Db	73	GAGCAGATGGGCACTCAAGGGTTTGAAGAACTGCTGGGGGACATTCGGGCCCAAGGCATG	132
QY	139	AAGGAGCAGAAAGTTCGAGAGCTACTTCGGCCCGCAAAATCGCGCTCGAGCGCGCATGAGC	198
Db	133	AAGGAGCAGAAAGTTCGAGAGCTACTTCGGCCCGCAAAATCGCGCTCGAGCGCATGAGC	192
QY	199	ATATACCAATTCCTGATTGTAGTTGGAGAGACAGGCATGGAACTCTCACAAATGAAGCT	258
Db	193	ATATACCAATTCCTGATTGTAGTTGGAGAGACAGGCATGGAACTCTCACAAATGAAGCT	252
QY	259	GGTAACTCACTAGTATATTGCAAGGAATGTTCACACCGGACAAATAGATTACTGGAAGCG	318
Db	253	GGTAACTCACTAGTATATTGCAAGGAATGTTCACACCGGACAAATAGATTACTGGAAGCG	312
QY	319	GGATCAACGCACTTTATGTTTATATGSCAAGCCTCCTGATATGAGAAACAAGACCTT	378
Db	313	GGATCAACGCACTTTATGTTTATATGSCAAGCCTCCTGATATGAGAAACAAGACCTT	372
QY	379	GCTAAAGATACTCAAAAAGATATGATGCAACCAAGATCTGACTGAGGCGAGTAGAGTA	438
Db	373	GCTAAAGATACTCAAAAAGATATGATGCAACCAAGATCTGACTGAGGCGAGTAGAGTA	432
QY	439	GGAGATTAACATCGGANTGGAATAATGACACAGAGACGTGAAGGTCACACAGGCACAC	498
Db	433	GGAGATTAACATCGGANTGGAATAATGACACAGAGACGTGAAGGTCACACAGGCACAC	492
QY	499	AACCAAGATTGTAACCGGCTATTAAAGACTTATGGGGGTTCCGTGTTGAGAGCACCTTC	558
Db	493	AACCAAGATTGTAACCGGCTATTAAAGACTTATGGGGGTTCCGTGTTGAGAGCACCTTC	552
QY	559	GAAACAGACAGAGATATGTCAGGCCCTTCGATTAACAGTAAGAGGTTGCGTGTGCTCA	618
Db	553	GAAACAGAGAGAGATATGTCAGGCCCTTCGATTAACAGTAAGAGGTTGCGTGTGCTCA	612
QY	619	GAAATATGAGACTCCCTTACTTTTGGGGCTCCACAGGTTCCCTGTCATTTAATGATCCA	678
Db	613	GAAATATGAGACTCCCTTACTTTTGGGGCTCCACAGGTTCCCTGTCATTTAATGATCCA	672
QY	679	AGTTCCAAAGAAATACCTGTGATGAGAAATTTGATGCCAAGGTTTGGAGACTTGAA	738
Db	673	AGTTCCAAAGAAATACCTGTGATGAGAAATTTGATGCCAAGGTTTGGAGACTTGAA	732
QY	739	CTCACACATGACAGCTTCATTTGATTCGATCCTGTGGATGTGATATTTGATAGC	798
Db	733	CTCACACATGACAGCTTCATTTGATTTGTGCATCTGTGGATGTGATATTTGATAGC	792
QY	799	ATCAAAAGTATCGGGGGGCAAAAGCTCTGCAAACTTATTCGCAACATGGGTCATAGAA	858
Db	793	ATCAAAAGTATCGGGGGGCAAAAGCTCTGCAAACTTATTCGCAACATGGGTCATAGAA	852

1141 AAAAAAAAAAAGGCTGGTGGTAAGAAAGAAATATTTGGATGCTGTGATGTAACCTA 1200
QY 1249 CGAGTCAGCAAGCAGCGGTGGCTGATCATTGCTTATTTACTTCCCTGTTTAA 1308
Db 1201 CGAGTCAGCAAGCAGCGGTGGCTGATCATTGCTTATTTACTTCCCTGTTTAA 1260
QY 1309 ACTCAGAGCTTTGGTAAAGTTTGCATGTTTCAAGCTGGGGTAAAGTTAGTTGTTG 1368
Db 1261 ACTCAGAGCTTTGGTAAAGTTTGCATGTTTCAAGCTGGGGTAAAGTTAGTTGTTG 1320
QY 1369 AAGAGTTGTTGATCAGCAATTAACCAACTTATCGCTGTTT 1411
Db 1321 AAGAGTTGTTGATCAGCAATTAACCAACTTATCGCTGTTT 1363

RESULT 5
US-08-455-968E-2
; Sequence 2, Application US/08455968E
; Patent No. 5874283

GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-455-968E-2

Query Match 20.7%; Score 302.6; DB 2; Length 1144;
Best Local Similarity 56.6%; Pred. No. 5.4e-70;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 85 ATGGGATCAAGGCTTGGCAAACTGCTGGCGCAATGCGCCCAAGCGCATAGAGAG 144
Db 1 ATGGGATCAAGGCTTGGCAAACTGCTGGCGCAATGCGCCCAAGCGCATAGAGAG 60
QY 145 CAGAAGTTGAGAGCTACTTGGCGCAAAATCGCGTCAGCGCAGCAGCATATAC 204
Db 61 AATGACATCAAGAGCTACTTGGCGTAAGTGGCCATGATGCTCATAGAGATTAT 120
QY 205 CAGTCTCTGATTTAGTTGGAAGAGCAGCATGAAACTTCACAAATGAAGCTGGTGA 264
Db 121 CAGTCTCTGATTTAGTTGGAAGAGCAGCATGAAACTTCACAAATGAAGAGGAGTGA 177

QY 265 GTCACTACTCATTTGCAAGGATTTCAACCGGACATTAAGATTACTGGAACGGGAATC 324
Db 178 ACCACACACCCCTGATGAGGATATTTCTACGACCAATTCGATGCAAGAGGACATC 237
QY 325 AAGCAGTTTATGTTTATGATGGCAACCTCTGATATGAAAGCAAGAGCTTGTAA 384
Db 228 AAGCCCTGTATGCTTTGATGGCAACCCGCCACAGCTCAAGTCAGGCGAGTGGCAAA 297
QY 365 AGATCTCAAAAAGAGATGATGCAACCAAGATCTGACTAGGACATAGAGTAGAGAT 444
Db 298 CGACATGAGCGCGGCTGAGGAGAGACAGTACGAGCGGCTGAGCTGTGGGGCC 357
QY 445 AAGATGCGATTTGAAAAATTCAGCAAGAGAGAGCTGTAAGGTCAAGAGCAACAGAA 504
Db 358 GAGCGAGAGTGGAAAAATTCATAAGCGGCTGTGAAGGTCACTAAGCAGCACAATGAT 417
QY 505 GATTGTAAAGCGCTTATTAAGCTTAAAGGAGTTCCTGTTGATAGAGCAGCTTGTGAAGA 564
Db 418 GAGTCANACATCTGCTGAGCTGATGGCATCCCTTATCTTATGATGACACAGTAGGCA 477
QY 565 GAGCGAATGTGAGCCCTTTGCAATAAGCAATAGGTGCTGCTTCAGAAAT 624
Db 478 GAGCGCAGCTGTGCTGCTGCTGTAAGGCTGCAAGATCTATGCTGCGCTACCGAGCAG 537
QY 625 ATGACCTGCTTACTTTGGGGCTCCAGCGTCTGCTGCTGCTTAAATGATCCAACTTCC 684
Db 538 ATGACCTGCTTACTTTGGGGCTCCAGCGTCTGCTGCTGCTTAAATGATCCAACTTCC 597
QY 685 AAGAAATACCTGTGATGATGATTTGATGTTGCCAAGGTTTGGAGAGCTTGAATCACC 744
Db 598 AAAAAGCTGCAATTCACCAATTCACCTGAGCGGATTCGACGAGAGGCGGCTGAAC 657
QY 745 ATGACCACTTCATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 804
Db 658 CAGGAACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
QY 805 GGTATCGGGGCAACAGCTGTAATTTGCTGATGATGATGATGATGATGATGATGATGAT 864
Db 718 GGTATCGGGGCAACAGCTGTAATTTGCTGATGATGATGATGATGATGATGATGATGAT 777
QY 865 TTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 924
Db 778 GTGGGGGCACTTGAACCCCAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 837
QY 925 CGACCTTGTTCAGAGCCTTAATG---TCACATTTGATATTTCTGAGCTTAATGAGCT 981
Db 838 CACCACTCTCTTGTGAACCTGAGGTGCTGAGCCCAAGATGCTGTGAGAGTGAAGTGAGC 897
QY 982 GCACCTGATGAGAGGCTCTCATAGTTCTGCTGTAAGATGATGATGATGATGATGATGAT 1041
Db 898 GAGCAATGTAAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
QY 1042 CGGTTGCAAAAGCCATGAGAGATCAATCTGCCAAGATTAATCTGCGCAGAGAGAA 1101
Db 958 CGAATCCGCAAGTGGGTCAAGAGGCTGTAAGAGCCGCAAGAGCAGCAGCAGGCGCC 1017
QY 1102 CTCGAGTCTTTTCAAGCAATGCGCACCATCAGCAGC 1142
Db 1018 CTCGATGATTTCTTCAAGCTACCGGCTCACTCTTCAAGC 1058

RESULT 6
US-08-455-968E-28
; Sequence 28, Application US/08455968E
; Patent No. 5874283
GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:

```

; ADDRESSER: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (polynucleotide)
; US-08-455-968E-28

```

```

Query Match      20.7%   Score 302.6; DB 2; Length 1144;
Best Local Similarity 56.6%; Pred No. 5.4e-70;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

```

```

QY 85 ATGGGATCAAGGGTTTGACGAATGCTGGCGACATGCGCCCAAGCGATGAAGAG 144
   |||||  ||  ||  |||||  ||  ||  |||||  ||  ||  |||||
DB 1 ATGGGATTAAGAGCGCTGGCAAACTAATGTGATGTGCGCCCAAGCGCATCGGAG 60
QY 145 CAGAAGTTCAGAGCTACTTCGGCGCAAAATCGCGGTGAGCGACGATGATATAC 204
   |||||  ||  ||  |||||  |||||  |||||  |||||  |||||  |||||
DB 61 AATGACATCAAGAGCTACTTGGCCCTTAAGGTGGCCATGATGCTCTATGAGCATTTAT 120
QY 205 CAGTTCTGATTTAGTTGGAAGAGCAGCATGAAATCTCAACAATGAAGCTGTGAA 264
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 121 CAGTTCTGATTTAGTTGGAAGAGCAGCATGAAATCTCAACAATGAAGAGCTGTGAG 177
QY 265 GTCCTGATTTGCAAGGATGTTCAACCGGCAATTAAGATTAAGTCTGGAACCGGAGATC 324
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 178 ACCACGACCACTGATGGCATGTTCTACCGCACCATTCCATGATGAGAGAGGCGATC 237
QY 325 AAGCAGTTATGTTTGTGATGCAAGCTCTGATATGAAGAAACAGAGCTTGCTAA 384
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 238 AAGCCGCTGATGCTTGTGATGCAAGCGCGCACAGCTCAAGTCAGCGCGTGCACAA 297
QY 385 AGATATCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCGATGAGTAGAGAT 444
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 298 CGCAGTAGCGGGGGGTGAGGCAAGAAAGACAGCTCAGACAGCTCAGCTCTGSGGCG 357
QY 445 AAAGATCGATGAAAAATTGAGCAAGAGAGCTGTAAGGTGACAGAGCAGACAGCAA 504
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 358 GAGCAGAGAGGTGAAAAATTCACTAAGCGCTGTAAGCGCTCACTAAGCAGCATATAT 417
QY 505 GATGTAAAGCGCATTAAGCACTTAATGAGGGTCCCTGTTGAGAGCAGCTTCAAGCA 564
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 418 GAGTGCAAAATCTGCTGAGACCTCATAGGCGATCCCTTATCTGATGACCCAGTAGAGGA 477
QY 565 GAAGCAGATGTGACGCGCTTTCATAAAGATGAAGGTGCTGCTGCTGCTGCTGCTGCTGCT 624
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 478 GAGCGCAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
QY 625 ATGAGCTCCCTTACTTTGGGGCTCCAGGTTCTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

DB 538 ATGAGCTGCTTCACCTTGGCGACGCCCTGTGCTAATGCGACACCTGACTGCGAGTGAAGCC 597
QY 685 AAGAAAATACCTGTATGAGATTTGATCTTCCCAAGGTTTGGAGAGCTTGAACCTAC 744
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 598 AAAAGCTGCGCAATCCAGAAATTCACCTGAGCGGATTTGCTGAGGCGGTGTAAC 657
QY 745 ATGAGCAGCTTATGATTTGATGCTGATCTGATGATGATGATGATGATGATGATGATGATGAT 804
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 658 CAGAAACATTTGTGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 805 GGTATGGGGGCGCAACAGCTGTGAATTTATGCTCAACATGSGTCCATGAAAGATC 864
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 718 GGTATGGGGGCGCAACAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 777
QY 865 TTGAGAACTTTAATTAAGACATATCAAAATTCCTGAGACTGCGCTTACCAAGAGCT 924
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 778 GTGCGGCGACTTGACCCCAACAGATACCTTGCCAGAAATTTGGCTCCACAGAGAGCT 837
QY 925 CGACGCTTGTCAAGAGAGCTTAATG---TCAATTTGATTAATCTGAGCTAAATGACT 981
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 838 CACACAGCTCTTCTTGTGAACCTGAGGTGCTGAGCCAGAGTCTGTGAGCTGTAAGTGAAGC 897
QY 982 GCACCTGATGAGAGAGGTCTCATAGTTTCCGTGAAGAAATTAATGTTCAACAGAGAT 1041
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 898 GAGCCAAATGAAGAGAGCTGATCAAGTTCAATGCTGTGTAAGAGAGCTTCTGAGAGAG 957
QY 1042 CGGGTACAAAGGCCATGAGAGATCAATCTGCCCAAGATTAATGCTGCCAAGAGAGA 1101
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 958 CGAATCCGAGTGGGTGCAAGAGGCTGAGTAAGAGCCGCCAAGCAGAGCAGCAGAGGCGCG 1017
QY 1102 CTCGAGTCCCTTTTCAAGCCATGCGCCACACATCAGCACC 1142
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1018 CTCGATGATTTCTTCAAGGTGACGCGCTCAGCTCTTGAAGC 1058

```

```

RESULT 7
US-08-455-968E-9
; Sequence 9, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; TITLE OF INVENTION: Mamalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 104..1237
 ; US-08-455-968E-9

Query Match 18.9%; Score 276.6; DB 2; Length 2033;
 Best Local Similarity 55.9%; Pred. No. 4.5e-63;
 Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 85 ATGGGCATCAAGGGTTTGACAAACTGCTGGGACAAATGCGCCCAAGCGCATGAAGAG 144
 DB 104 ATGGAATAATTCAGCGCCCTTGCCAAACTAATGATGATGGCCCCCAGAGGCCATCGTAG 163
 QY 145 CAGAAGTTGAGAGCTCTTCGGCGCAAAATGCGCCGACGATGAGCATGATATAC 204
 DB 164 AATGACATCAAGAGCTACTTGTGCGCAAGTGCCATGAGCTCCATGAGCATCTAC 223
 QY 205 CAGTTCTGATTTGATTTGAGAGCAGCAGCATGAAACTCTCAAAATGAAGCTGGTGA 264
 DB 224 CAGTTCTGATTTGATTTGAGAGCAGCAGCATGAAACTCTCAAAATGAAGCTGGTGA 280
 QY 265 GTCACATGCTATTTGACAAAGTGTTCACACCGGACATTAAGATTCTGGAAGCGGAATC 324
 DB 281 ACCACAGCC---TGATGGGATGTTCTACCGTACCATCGC---CATGAGAAATGGCATC 334
 QY 325 AAGCCAGTTATGTTTGTGATGAGCAAGCTCTGATATGAGAAACAGAGCTTCTAAA 384
 DB 335 AAGCCTGTGACGCTCTTGTGAGCAAGCTCTGATATGAGAAACAGAGCTTCTAAA 394
 QY 385 AGATACCTCAAAAAGAGATGATGACCAAAAGATCTGACGAGCGATGAGGTAGAGAT 444
 DB 395 CGCAGTGAAGGGCGCGCGGAGGCTGAGAGCACTGACAGAGGCTCAGAGCGTGGATG 454
 QY 445 AAAGTGCATTTGAAAAATTGACCAAGAGAGCTGTAAAGTCAACAAGCAACACAGAA 504
 DB 455 GAGAGGAGGGGTGAGAACTTACCAAGAGGCTGTGAAGTCAACAAGCAACACATAT 514
 QY 505 GATTGTAAGCGGCTATTAGACTTATGGGGGCTCTGTTGAGAGCACTTCTGAGCA 564
 DB 515 GAGTGAACAACACTGCTGACGCTCATAGGGGATCTTACTTGATGACACCGAGGCA 574
 QY 565 GAAGCAAGATGTCGAGGCTTGTGCAATAAGATAGTGTGCGTTCCTCAAAAT 624
 DB 575 GAGGCAAGCTGTGCTGCGCTGGCAAGAGCTGCAAAAGTCTATGCTGCGGCCA 634
 QY 625 ATGACCTCCCTTACTTTGGGGCTCCACGCTTCTCTCATTTAATGATCCAGATTCC 684
 DB 635 ATGACCTCCCTTACTTTGGGGCTCCACGCTTCTCTCATTTAATGATCCAGATTCC 694
 QY 685 AAGAAATACCTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGAT 744
 DB 695 AAGAAATACCTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGAT 754
 QY 745 ATGACCAAGTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 804
 DB 755 CAGAGGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 814
 QY 805 GGTATCGGGGGGCAACAGCTGGAACCTAATCGCAACATGCGGCTATAGAAAGCATC 864
 DB 815 GGCATTTGGCGGCAAGCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 874
 QY 865 TTGGAGATCTTAAATAAGACAGATATCAATTTCTGAGAGCTGGCTTACCAAGAGCT 924
 DB 875 GTGAGGGCGGTGAGACCCAGCAAGTACCCGTTCCAGAAAGAGGCTCCACAGAGAGCC 934
 QY 925 CGAGCTTGTTCAGAGAGCTTAA---TGTCACATTTGATATTTCTGATGATTAATGAGACT 981
 DB 935 CAGCAGCTCTTCTGAGAGCAGAAAGTAGTGAGCCAGAGTCTGTGAGTGAAGTAGAGC 994
 QY 982 GCACCTGATGAGAGGCTCATTAAGTTTCTGTGTAAGAGATTAAGTTTCAACAGAGAT 1041

DB 995 GAGCCAAATGAAGAAGAGCTGTCTCAAAATTTATGTTGCTGAAGAGCAGTTTTCAGAG 1054
 QY 1042 CGGGTGACAAAGCCCATAGAGAGATCAAAATCTGCGCAAGATAATGTCGCAAGAGAGA 1101
 DB 1055 CGAATTCGAGAGTGGGCTCAACGGCTGTGTAAGAGCCGCGAGGAGCAGACCCAGGAGCC 1114
 QY 1102 CTCGAGTCTCTTTTCAGCCCAACTGCGCCACATCAGCAGCCGCTAAAGAGAGAGACT 1161
 DB 1115 CTCGATGATTTCTTAAAGGTGACAGGCTCACTCTCTCAGC-----TAAGCCG 1162
 QY 1162 TCGATTAACAGCAGCAGCTGCGGACAGAAACAAAGGCTGTGGAAGAAGAAA 1221
 DB 1163 AAGGAGCCGCAACCCAGAGGCGCTGTGAAGAAAGAAAGACTGGGAGCGCGGAGAG 1222
 QY 1222 T 1222
 DB 1223 T 1223

RESULT 8
 US-08-455-968E-4
 ; Sequence 4, Application US/08455968E
 ; Patent No. 5874283
 ; GENERAL INFORMATION:
 ; APPLICANT: Harrington, John L.
 ; APPLICANT: Isieh, Chih-Lin
 ; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, 8th Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/455,968E
 ; FILING DATE: 30-MAY-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Apple, Randolph T.
 ; REGISTRATION NUMBER: 36,429
 ; REFERENCE/DOCKET NUMBER: 18985-000100
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-576-0300
 ; TELEFAX: 415-576-0300
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1930 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; US-08-455-968E-4

Query Match 18.6%; Score 272.2; DB 2; Length 1930;
 Best Local Similarity 55.1%; Pred. No. 6.2e-62;
 Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;

QY 85 ATGGCATCAAGGGTTTGACAAACTGCTGGCGGACAAATGCGCCCAAGCGCATGAAGAGAG 144
 DB 1 ATGGAATTCAGCGCCCTTGCCAAACTAATGATGATGGCCCCCAGAGGCCATCGTAG 60
 QY 145 CAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 204
 DB 61 AATGACATCAAGAGCTACTTGTGCTGTAAGTGGCATGAGTGTCTCATGAGCATCTAC 120

```

QY 205 CAGTTCCTGATGTTAGTGGAAAGACAGGCAATGGAACCTCTCAACAATGAAGCTGGTAA 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CAGTTCCTGATGTTAGTGGAAAGACAGGCAATGGAACCTCTCAACAATGAAGCTGGTAA 177
QY 265 GTCACTAGTCAATTTGGCAAGGATGTTCAACCGGACCAATAGATTAAGTACGAAAGCGGCAATC 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 ACACACAGGCTGATGGGCAATGTTAT-----GGCAACCACTCCGCAATGGCAATGGCATC 231
QY 325 AAGCCAGTTTATGTTTATGAGCAAGCTCTCTGATATGAAGAAACAAGCTTCTAA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 AAGCTGTGTAGCTTTTATGAGCAAGCTCTCTGATATGAAGAAACAAGCTTCTAA 291
QY 385 AGATTAATCAAAAAAGATGATGCAACCAAGATGACTGAGGAGTATGAGTATGAGAT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 CGCAGTGAAGAGCGCGCGCAGGCTGAGAAAGCAACTGACGAGGCTGAGAGGCTGGGATG 351
QY 445 AAGATGCGATTTAAATAATTTAGCAAGAGGACTGTAAGGTCACAAAGGCAACACGAA 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 GAGGAGAGGTGAGAAAGTTACCAAGAGGCTGGAAGGTCACCAAGCAACCAATGAT 411
QY 505 GATTGTAAGGCTATTAGACTTATGAGGAGTCTCTGTTAGAGCGCACTTCTGAAGA 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 GAGTGAACACCTGCTGAGCCTCAATGGGCACTCTTACCTTGATGACCCAGAGAGCA 471
QY 565 GAAGCAAAATGTGCACCTTTGCAATAACGATAGGTTGCTGCTTCTGAGAAAT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 GAGGCAAGCTGTCTGCTGCAAGGCTGCAAAAGCTATGCTGCGCGCAGGAGATC 531
QY 625 ATGAGCTCCCTTATTTGGGCTCCAGGTTCTCTCTCAATTTATGATTCGAAGTTCC 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 ATGAGCTCCCTTATTTGGGCTCCAGGTTCTCTCTCAATTTATGATTCGAAGTTCC 591
QY 685 AAGAAATACCTGTGATGAAATTTGATGCGCAAGTTTGGAGGAGCTTGAAGTAC 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 AAGAGCTGCTCCATCCCAAGATTCATCTGAGCCGCTGCGAGAGAGCTGGCTTAA 651
QY 745 ATGAGCAATCTTATTTGATGATCTCTGATGATGATGATGATGATGATGATGAT 804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 CAGGAGAGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
QY 805 GATATGCGGCGGCAACACTGCAAACTATGCTCAACATGGTGCTCAATGAAGATC 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 GGCATTTGCGCCCAAGCGGCTGTGATCTCATCCAGAAACATTAAGAGATGAGAGATC 771
QY 865 TTGAGAAATCTTATTAAGACAGATATCAATTTCTGAGGAGCTGGCTTACCAAGAGCT 924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 GTGAGCGGCTGAGACCCAGCAAGTACCCCTTCCAGAGAACTGCTCCACAAAGAGCC 831
QY 925 CGAGCGTGTTCAGAGAGGCTAA---TGTCAATGATGATGATGATGATGATGATGAT 981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 CAGTACGCTTCTCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 891
QY 982 GCACCTGATGAGAGGCTCTCATTAAGTTTCTGTTAAAGATTAATGTTTCAACGAAGAT 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 GAGCAATATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
QY 1042 CGGAGTCAAGAGGCAATGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 952 CGAATTTGCGAGTGGGTCAAGGCGCTGATGATGATGATGATGATGATGATGATGATGAT 1011
QY 1102 CTGAGTCTTTTTTCAAGCAACTGCGACACATGAGCAAGCTTAAAGAGAGACT 1161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1012 CTGAGTCTTTTTTCAAGCAAGCTGCTCTCTCAGC-----TAAACGCG 1059
QY 1162 TCGATTAACAAAGCAAGGAGCTGCAACAAAGAAACAAAGGCTGTGTAAGAAAGAA 1221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 AAGGAGCCAAACCCAAAGGCGCTGCTAAGAAAGAAAGCAAAAGACTGAGGAGCGGAG 1119
QY 1222 T 1222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 T 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
US-08-455-968E-6
; Sequence 6, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-May-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-6

Query Match      18.5%; Score 270.2; DB 2; Length 1149;
Best Local Similarity 55.5%; Pred. No. 1,7e-61;
Matches 599; Conservative 0; Mismatches 453; Indels 27; Gaps 3;

QY 85 ATGGGCAATCAAGGCTTATGAGCAAACTGCTGCGGCAATGCGCCGAGAGATGAGATATAC 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGGGATTTAAAGGTTTGAATGCAATTAATGGAACATGTTCCCTGCTATCAGAGAA 60
QY 145 CAGAAATGCAAGAGCTACTTGGCGCCCAAAATCGCCGCGAGCCAGATGAGATATAC 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AGCGATTTCAAGAGCTTTTGGCAGAAAGTTGCGATGATGCTCTATCTCTATAT 120
QY 205 CAGTTCCTGATTTAGTTGGAAGGACGACATGAAACTCTCAAAATGAAGCTGTGAA 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CAGTTTATTAATGCTGTGAAGCAGCAAGCGGTGGGCAAGTTGACCAATGAAGCGGTGAA 180
QY 265 GTCACTAGTATTTGCAAGGAATGTTCAACGCGACAAATGAATTAATGGAAGCGGATC 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 ACAACGTACACTGATGAGGTATGTTTATAGGACACTGAGAAATGATGATTAACGGATATC 240
QY 325 AAGCAGTTTATGTTTATGATGAGCAAGCTCTCTGATATGAAGAAACAAGAGCTGTAA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 AAGCTTTGATGCTTCGACGCAAACTCCAGCTTGAATCTCAAGATTTACAAAG 300
QY 385 AGATTAACAAAGAGATGATGCAACCAAGATCTGACTGAGGCGAGTGTAGATGAGAT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 CGGCTTCAAGAGGAGGAGCAAGAAATAAGTGCAGAGGACACACAGAAATTGAA 360
QY 445 AAGATGCGATTTGAAAAATTGAGCAAGAGAGCTGAAGGTGCACAGGCAACACAGAA 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AAGATGAAGCAAGAAAG-----AGATTGTGAAGGTCTCAAAAGAGCATATGAA 411

```

```

QY 505 GATTGTAACGGCTATTAGACTTATGGGGTCTCTGTTGAGAGCACTTCTGAAGCA 564
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 412 GAAGCCCAAAATTTACTAGACTAATGGAAATCCATATTAATAGGCCCAACGAAAGCT 471
QY 565 GAAGCAGAAATGTCAGCCCTTTCATAAACGATAGTAGTTCGCTGTTGCTTCAGAGAT 624
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 472 GAGGCTCAATGCTGAGTGTGCAAGAAAGGAAAGGTGTATGCGCAGCAAGTGAAGAT 531
QY 625 ATGAGTCCCTTACTTTTGGGGCTCCAGGGTTCCTGCTCATTTAATGATCCAGTTCC 684
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 532 ATGACACACTCTGTTATGAACACCCCTTCTGTTGAACATTTGACTTTTTCAGAGCC 591
QY 685 AAGAAATACCTGTGATGGAATTTGATGTCGAAGTTTTCAGAGAGCTTGAATCACC 744
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 592 AAGAAGAACCCGATTCAGAAATAGTACTGAATAGTTTGAAGAGACTGACTTGACA 651
QY 745 ATGAGACCATTTCTATTTATTTTGTGATCCCTGTGTGATGTGACTATTTGATAGCATCAA 804
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 652 ATGAGACAGTTTGTGATCTTTGCAATATGCTGTTGTTGATGCTGTAAGAGCATAGA 711
QY 805 GGTATCGGGGGGCAACAGCTCTGAACTTATTCGTCACATGGGTCCATGAAGCATC 864
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 712 GGTGTTGTCACAGTACAGCCCTTAATATGATAAAGCAATGGATCCATGAGAAATATC 771
QY 865 TTGGAGAACTT-----ATAAGACATATCAAAATTCCTGAGAGACTCG 909
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 772 GTGAGATTATTTGAATCTGGGAGTCACAACTAATGGAATCCCGAAGAGCTGG 831
QY 910 CTTTACCAAGAAAGCTGAGCTTGTTCAGAGAGCTTAAATGTCACATTGGATA--TTCT 966
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 832 CTTTACCAAGAAAGCTGAGCTTGTTCAGAGAGCTTAAATGTCACATTGGATA--TTCT 966
QY 967 GAGCTAAATGAGCTGAGCTGATGAGAGAGCTGATGAGAGAGCTGATGAGAGAGAT 1026
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 892 AACTGAAATGTCGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951
QY 1027 GATTTCACAGAGATCGGTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 952 AAATTCAGTGAAGAAAGAGATCTGATATATCAATGAGAGAGAGAGAGAGAGAG 1011
QY 1087 TCGTCGCAAGAGAGAGCTGAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1145
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 1012 GGCATTCAGGATGATTAGATGAGGTTCTTCAGAGAGAGAGAGAGAGAGAGAGAG 1070

RESULT 10
US-08-757-653-175
; Sequence 175, Application US/08757653
; Patent No. 5843669
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,653
; FILING DATE:
; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 175:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1020
; US-08-757-653-175

Query Match 10.9%; Score 159.6; DB 2; Length 1023;
Best Local Similarity 54.6%; Pred. No. 1.5e-32;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAGAGGAGAGAGTTGAGAGAGTCTGCGCGCAAAATCGCGAGCCAGCA 193
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 29 CAAGAAAGAAATTTAGCTTACAAAACCTATACGGGAAAAAATTCGATGACGCTCTTA 88
QY 194 TGAGCATATACCAGTCTTCTGATTTGAGTGAAGAGACAGCATGGAATCTGCATAATG 253
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 89 ATGCATCTACCAATTTTGTCCAAATTAAGACAGAAAGATGGAATCTGCTATGATGATT 148
QY 254 AAGCTGTGGAAGTCTGCTGATTTGCAAGAGATGTTCAACCGGACATTAAGATTACTGG 313
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 149 CAAGAGTAGATATATCCCTCCACCTTAAGCGGCTCTTTACAGACATTAATACCTAATG 208
QY 314 AAGCGGAGATCAAGCACTTTATTTTATGAGCAAGCTCTCTGATATGAGAAACAAG 373
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 209 AAGCTGTGATTAACCTGTGATTTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 268
QY 374 ACCCTGTAAAGATATCTCAAAAGAGATGATGCAACCAAGATCTGAGAGAGAGAG 433
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 269 AACTGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328
QY 434 AGGTAGAGATTAACATCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 493
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 329 AAAAAGGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
QY 494 AACACAGAGAGATTTGTAACGAGCTATTAAGACTTATGAGGGGTTCTGTTGAGAGCAG 553
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 389 TCGTATCGAGAGATGCAAAAACCTTAGAGCTTATGAGGATTCCTATGTTCAAGCAG 448
QY 554 CTTTGAAGCAGACAGCAAAATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 613
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 449 CTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
QY 614 CTTGCAAGATATGAGAGTCCCTTACTTTTGGGCTCCAGAGAGAGAGAGAGAGAGAG 673
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 509 CTAGTCACATTAATGATTCCTGACTTTTGGAGCTCCAGAGAGAGAGAGAGAGAGAG 568
QY 674 ATCCAGATTCAGAGAAATACCTGTGATGAGATTTGATGTTG 715
D 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
D 569 TAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610

RESULT 11
US-08-823-516-78
; Sequence 78, Application US/08823516
; Patent No. 5984069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.

```


SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: 1..1020
US-08-758-314-114

Query Match 10.9%; Score 159.6; DB 3; Length 1023;
Best Local Similarity 54.6%; Pred. No. 1.5e-32;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAGAGCAGACAGATGAGAGCTACTTGGCGCGCAAAATCGCGCTGACGCGCAGCA 193
DB 29 CAGAAAAGAAATGAGTTAGAAAACCTATACGGAAAAAATCGCAATCGACGCTCTTA 88
QY 194 TGAGCATATACCACTTCTGATTGTAGTTGGAAGACAGCATGAACTCTACAAATG 253
DB 89 ATGCATCTACCAATTTTTCACAAATAGACAGAAAGTGAACCTCCACTTATGATT 148
QY 254 AAGCTGAGAGTCACTAGTATTTTGGACGAATGTTCAACCGACAAATAGATTACTG 313
DB 149 CAAAGGTAAGATTAACCTCCACCTAGCGGCTCTTTTACAGACAAATTAACCTAATG 208
QY 314 AAGCGGGAATCAAGCAGTTATTTTATGTCGAAGCTCTCTGATATGAAGAAACAG 373
DB 209 AGGCTGGAATTAACCTGTGTATTTTATGTAAGAACTCCGAATTCAAAAGAAAG 268
QY 374 AGCTTGCTAAAGATTAAGTCTCAAAAAGAGATGCAACCAAGATCTGATGAGCAGT 433
DB 269 AGCTGGAAGAAAGAAAGAACGAGAGAGAGCTGAAGAAAGTGAAGAGACGCTG 328
QY 434 AGGTGAGATTAAGATGACATTAATAATGAGCAAGAGACTGTAAGCTCACAGGC 493
DB 329 AAAAAGAGATAGAGAGAGCAAGAAATATGCCCCAAGAGCAACAGGTAATGAAA 388
QY 494 AACACACGAAGATTTAAACGCTATTAAAGCTTATGCGGGTTCCTGTGTAAGAGC 553
DB 389 TGCTCATGAGAGATGCAAAAACCTTAAGAGCTTATGGAATTCCTAATAGTTCAAGC 448
QY 554 CTTTGAAGCAGACGAAGATGTCACGCCCTTTCATTAACGATTAAGTGGCTGTTG 613
DB 449 CTAGGAGAGAGAGAGCCCAAGCTCATATATGCGCGAAAGGAGACGTGTATGCTGG 508
QY 614 CTTCAAGAAATATGACTCCCTTACTTATGCGGCTCCACGGTTCCTTCATTTAATG 673
DB 509 CTAGTCAAGATTAAGATTCCTACTTTTGGAGCTCCAAACCTGTTAAGAACTTAACA 568
QY 674 ATCCAAATTCAGAAATACCTGTGATGGAATTTGATGTTG 715
DB 569 TAACGAGAAAAAGAAAGTTGCTGCGAAAAATGTCTACGTCG 610

RESULT 13
US-08-758-314-114
Sequence 114, Application US/08758314
Patent No. 6090606

GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamlchev, Victor I.
APPLICANT: Lyamlchev, Natasha
TITLE OF INVENTION: Improved Cleavage Agents
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,314
FILING DATE: 02-DEC-1996

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FORS-02575
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ. ID NO. 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1020
US-08-758-314-114

Query Match 10.9%; Score 159.6; DB 3; Length 1023;
Best Local Similarity 54.6%; Pred. No. 1.5e-32;
Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 134 CGATGAGAGCAGACAGATGAGAGCTACTTGGCGCGCAAAATCGCGCTGACGCGCAGCA 193
DB 29 CAGAAAAGAAATGAGTTAGAAAACCTATACGGAAAAAATCGCAATCGACGCTCTTA 88
QY 194 TGAGCATATACCACTTCTGATTGTAGTTGGAAGACAGCATGAACTCTACAAATG 253
DB 89 ATGCATCTACCAATTTTTCACAAATAGACAGAAAGTGAACCTCCACTTATGATT 148
QY 254 AAGCTGAGAGTCACTAGTATTTTGGACGAATGTTCAACCGACAAATAGATTACTG 313
DB 149 CAAAGGTAAGATTAACCTCCACCTAGCGGCTCTTTTACAGACAAATTAACCTAATG 208
QY 314 AAGCGGGAATCAAGCAGTTATTTTATGTCGAAGCTCTCTGATATGAAGAAACAG 373
DB 209 AGGCTGGAATTAACCTGTGTATTTTATGTAAGAACTCCGAATTCAAAAGAAAG 268
QY 374 AGCTTGCTAAAGATTAAGTCTCAAAAAGAGATGCAACCAAGATCTGATGAGCAGT 433
DB 269 AGCTGGAAGAAAGAAAGAACGAGAGAGAGCTGAAGAAAGTGAAGAGACGCTG 328
QY 434 AGGTGAGATTAAGATGACATTAATAATGAGCAAGAGACTGTAAGCTCACAGGC 493
DB 329 AAAAAGAGATAGAGAGAGCAAGAAATATGCCCCAAGAGCAACAGGTAATGAAA 388
QY 494 AACACACGAAGATTTAAACGCTATTAAAGCTTATGCGGGTTCCTGTGTAAGAGC 553
DB 389 TGCTCATGAGAGATGCAAAAACCTTAAGAGCTTATGGAATTCCTAATAGTTCAAGC 448
QY 554 CTTTGAAGCAGACGAAGATGTCACGCCCTTTCATTAACGATTAAGTGGCTGTTG 613
DB 449 CTAGGAGAGAGAGAGCCCAAGCTCATATATGCGCGCAAAAGGAGAGCTGTATGCTGG 508

Tue May 27 16:02:04 2003

us-09-805-311-1.std.rni

Page 13

OY		CITCAAGAAATNTGACCTCCCTTACTTTGGGTGCAGCGTCCTCCTCATTAATGG	673
Dg	509	CGAGTCACAATTATATTCCTACTCTTGTCGCCAAGGTGAATTAGAATCTGCACC	568
OY	674	ATCCAAGTTCACAATAATFACCTGTGATGSAANTTGATSTTG	715
Dg	569	TTAGCGGAAAAAAGAAAGATGCGTSGGAAAAGATTTTACCTTG	610

Search completed: November 5, 2002, 11:02:54
Job time : 46.2144 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:28:52 : Search time 1251.78 Seconds
(without alignments)
15774.429 Million cell updates/sec

Title: US-09-805-311-1
Perfect score: 1463
Sequence: 1 cacgagatagctcgcgcgc.....aaaaaaaaaaaaaaaa 1463

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estcro:*
8: em_hic:*
9: gp_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778.8	53.2	901	10	BG837708
2	560.8	38.3	586	9	AI881599
3	536.4	36.7	550	10	BE639422
4	534.4	36.5	554	10	BE639421
5	517.8	35.4	553	9	AM562789
6	501	34.2	532	9	AM000375
7	465	31.8	648	9	AV913663
8	463.4	31.7	474	9	AM559173
9	462	30.6	467	9	AI861468
10	448	29.4	475	9	BE186786
11	429.8	28.7	733	9	AM562517
12	420.4	28.2	733	9	AM774700
13	413	28.1	553	9	AI834484
14	411.6	28.1	539	10	BG263233
15	402.2	27.5	414	9	AM288784
16	398	27.2	456	9	AI065689
17	378.8	25.9	632	9	AM585913

18	372.6	25.5	609	10	BG096545
19	370.4	25.3	754	10	BE416584
20	369.6	25.3	820	10	BG414505
21	364.6	24.9	645	9	AM691681
22	348	23.8	460	10	BE417817
23	347.2	23.7	683	9	AV918966
24	346.8	23.7	454	9	AV915644
25	341.8	23.4	363	9	AM562518
26	325.4	22.2	539	10	BF098320
27	317.2	21.7	470	10	AM288831
28	314.8	21.5	376	9	AI065346
29	314.8	21.5	579	9	AV833393
30	314.4	21.5	592	9	AI780966
31	308	21.1	526	10	BI419720
32	300.6	20.5	695	9	BE052579
33	273	18.7	549	9	AM696665
34	272.4	18.6	942	9	AL560007
35	262.2	18.1	903	9	AL531350
36	262.4	17.9	851	9	AL519300
37	248.6	17.0	689	10	BE311755
38	245.4	16.8	769	10	BG574950
39	244.4	16.7	660	10	BJ011893
40	242.8	16.6	806	9	AL560395
41	237.6	16.2	762	10	BG337603
42	234.6	16.0	866	10	BI827898
43	233.6	16.0	704	9	AM783795
44	233.4	16.0	963	12	CNS06WXR
45	233	15.9	785	10	BG756459

ALIGNMENTS

RESULT 1
LOCUS BG837708 901 bp mRNA linear EST 25-MAY-2001
DEFINITION Zm10_01f08_A Zm10_AAPC-ECORC_Fusarium_graminearum_corn_silk zea
mays CDNA clone zm10_01f08, mRNA sequence.

ACCESSION BG837708
VERSION BG837708.1 GI:44204031

KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE
AUTHORS Harris,L.J., Balcerzak,M., Allard,S., Sapano,A., Couroux,P., De Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A., Sprott D., and Finkler,N.A.
Expressed Sequence Tags from Maize Silk Six Hours After Silk Channel Inoculation with Fusarium graminearum
Unpublished (2001)
Contact: Harris, Linda J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1314
Fax: (613) 759-6566
Email: harris@em.agr.ca.

FEATURES

source
1..901
/organism="Zea mays"
/cultivar="CO388"
/db_xref="taxon:457"
/clone="Zm10_01f08"
/clone_id="Zm10_AAPC-ECORC_Fusarium_graminearum_corn_silk"
/tissue_type="Silk"
/dev_stage="4-5 days post-silk emergence"
/note="Vector: Bluescript SK(-)/XhoI-EcoRI; Site 1: EcoRI; Site 2: XhoI; Field-grown corn was silk channel-inoculated

Qy	Db
482 AGGTCCACAAAGGCAACCAACGAAGTTGTAACGGCTATTAGACTTATGCGGGCTTCG 541	435 CTGAGGAGTAGAGTAGAGATTAAGATGCGATTGAAAATTGAGCAAGAGCACTGTAA 494
495 AGGTCACAAAGGCAACCAACGAAGTTGTAACCACTATTAGACTATGCGGGCTTCG 554	
542 TTGTAGAGGACACTTCTGAAGCAGAGAGAA 573	
555 TTGTAGAGGACACTTCTGAAGCAGAGAGAA 586	

RESULT 3				
LOCUS	BE639422			
DEFINITION	BE639422	550 bp	mRNA	linear EST 30-Aug-2000
ACCESION	946033A02.Y2	946 -	tassel	primordium prepared by Schmidt lab Zea
VERSION	BE639422			
KEYWORDS	BE639422.1	GI:9952839		
SOURCE	EST.			
ORGANISM	Zea mays.			
	Zea mays.			

REFERENCE	1 (bases 1 to 550)
AUTHORS	Walbot,V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL COMMENT

Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 3227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.

```

FEATURES
Source
1. Location/Qualifiers
   1..550
   /organism="Zea mays"
   /cultivar="OH43"
   /db_xref="taxon:4577"
   /clone_lib="946 - tassel primordium prepared by Schmidt
   lab"
   /tissue_type="tassels"
   /dev_stage="just after the transition from vegetative to
   inflorescence development"
   /lab_host="XLOLR"
   /note="Organ: tassels; Vector: HybriZAP; Site:1: EcoRI;
   Site:2: XhoI; George Chuck dissected immature tassels
   between 1mm and 3mm. Sharon Stanfield prepared the cDNA
   library in HybriZAP. Sample insert size range was 350 bp
   to 3 kb with a 1 kb average."
BASE COUNT
177 a 116 c 149 g 108 t
ORIGIN

```

	Query Match	36.7%	Score 536.4	DB 10	length 550
	Best Local Similarity	99.6%	Pred. No. 1.66-66		
	Matches 548	Conservative 0	Mismatches 1	Indels 1	Gaps 1
QY	46	AGCGCGCCCGCCAGCCCGCACAGCCGCGCAGACGATGGCGCATCAAGGGTTTGACG	105		
Db	1	AGCGCGCGCCCGCCAGCCCGCACAGCCGCGCAGACGATGGCGCATCAAGGGTTTGACG	60		
QY	106	AAACTGCTGGCGGACAAATGCCGCCAAGGCGATGAAAGCAGCAGAAATTGCGAGACTACTTC	165		
Db	61	AAACTGCTGGCGGACAAATGCCGCCAAGGCGATGAAAGGAGCAGAAATTGCGAGACTACTTC	120		
QY	166	GGCGCGAAATGCGCGTCGAGCGCGCAGCATGAGCATATCCAGTTCCTGATTGTAGTTGGG	225		
Db	121	GGCGCGAAATGCGCGTCGAGCGCGCAGCATGAGCATATCCAGTTCCTGATTGTAGTTGGG	180		

QY	226	AGGCAAG-	-CATGAAACTCTCAAAATGAGCTGGTAAGTACTGTCATTGGCAAG	284
Db	181	AGCACTGTGTCATGGAACCTCTCACAATGAAAGCTGGTAACTACTAGTATTGGCAAG	240	
QY	285	AATGTTCAACCGGACAAATAGATTACTGGAAGCGGGAATCAAGCCAGTATGTTTTGA	344	
Db	241	AATGTTCAACCGGACAAATAGATTACTGGAAGCGGGAATCAAGCCAGTATGTTTTGA	300	
QY	345	TGGCAAGCCTCTCGATATGTAAGAAACAAGACTTGTCTTAAAGTACTCAAAAAGAAATGA	404	
Db	301	TGGCAACCTCTCGATATGTAAGAAACAAGACTTGTCTTAAAGTACTCAAAAAGAAATGA	360	
QY	405	TGCACCAAGATCTGACTGAGGCGAGTAGAGGTAGAGATTAAGATCCATTGAAAATT	464	
Db	361	TGCACCAAGATCTGACTGAGGCGAGTAGAGGTAGAGATTAAGATCCATTGAAAATT	420	
QY	465	GAGCAAGAGGACTGTAAAGGTCAACAAGGCACACACGACAGATTGTAACGGCTATTAAAG	524	
Db	421	GAGCAAGAGGACTGTAAAGGTCAACAAGGCACACACGAGATTGTAAAGGCTATTAAAG	480	
QY	525	ACTTATGGGGGTCTGTTGTAGAGGCAACCTTGAAAGCCAAAGCCAAAGTGGTACCCCT	584	
Db	481	ACTTATGGGGGTCTGTTGTAGAGGCAACCTTGAAAGCCAAAGATGTGACCCCT	540	
QY	585	TTGCATTAAC	594	
Db	541	TTGCATTAAC	550	

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
sermatophyta; Magnoliophyta: Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.		1 (bases 1 to 554)		
Walbot, V.		Maize ESTs from various cDNA libraries sequenced at Stanford University	Unpublished (1999)	Contact: Walbot V
		Department of Biological Sciences		
		Stanford University		
		855 California Ave,		
		Palo Alto, CA 94304, USA		
		Tel: 650 723 8227		
		Fax: 650 725 8221		
		Email: walbot@stanford.edu		
		Plate: 946033		row: A column: 02.
		Location/Qualifiers		

```

/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassels primordium prepared by Schmidt
lab"
/issue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="X10LR"
/note="Organ: tassels; Vector: HybridZP, Site 1: EcoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridZP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."
BASE COUNT
ORIGIN
145 a      131 c      110 g      168 t

```

Query Match 36.5%; Score 534.4; DB 10; Length 554;
Best Local Similarity 99.8%; Pred. No. 3e-68;
Matches 535; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 896 TTCCATGAGCTGCGCTTCACAAAGAGCTCGAGCGCTTGTTCAGAGAGCGCTTAATGTCACAT 955
|||||
Db 554 TTCCATGAGAGCTGCGCTTCACAAAGAGCTCGAGCGCTTGTTCAGAGAGCGCTTAATGTCACAT 495
|||||

QY 956 TGGATATTCCTGAGCTAAATATGACTGCACCTGATGAGAGGGCTCATTAAGTTCTCTGG 1015
|||||
Db 494 TGGATATTCCTGAGCTAAATATGAGCTGCACCTGATGAGAGGGCTCATTAAGTTCTCTGG 435
|||||

QY 1016 TAAAGATATATGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAATCAATCTG 1075
|||||
Db 434 TAAAGATATATGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAATCAATCTG 375
|||||

QY 1076 CCAGAAATTAATGCTGCAAGAGAGACTGAGTCTTTTTCAGAGCAAGCTCCACACAT 1135
|||||
Db 374 CCAGAAATTAATGCTGCAAGAGAGACTGAGTCTTTTTCAGAGCAAGCTCCACACAT 315
|||||

QY 1136 CACGACCGCTAAAGCGAGAGACTTCGATTAATAACAAAGCAGCAGCTCGAACAAGA 1195
|||||
Db 314 CACGACCGCTAAAGCGAGAGACTTCGATTAATAACAAAGCAGCAGCTCGAACAAGA 255
|||||

QY 1196 AAACAAAGCGCTGAGAAAGAAATTAATCTTGATGCTTGATGATACACTAGCAGTAC 1255
|||||
Db 254 AAACAAAGCGCTGAGAAAGAAATTAATCTTGATGCTTGATGATACACTAGCAGTAC 195
|||||

QY 1256 GAAAGCAGCGGTGCGTGATCACTTCGTTAGATTATTTAACTCCCTTTTAACTCGA 1315
|||||
Db 194 GAAAGCAGCGGTGCGTGATCACTTCGTTAGATTATTTAACTCCCTTTTAACTCGA 135
|||||

QY 1316 GCTTTGGTAAAGTTGCTCATGTTTCAAGCGGAGTAAAGTTAGTTGGTTTGAAGAGAT 1375
|||||
Db 134 GCTTTGGTAAAGTTGCTCATGTTTCAAGCGGAGTAAAGTTAGTTGGTTTGAAGAGAT 75
|||||

QY 1376 TGGGTACCAAGTAACAAACTATGCGTTTAACTCTGTGCTTTGAAGA 1431
|||||
Db 74 TGGGTACCAAGTAACAAACTATGCGTTTAACTCTGTGCTTTGAAGA 19
|||||

RESULT 5
AM562789 553 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.y1.660 - Mixed stages of anther and pollen zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AM562789
VERSION AM562789.1 GI:7216667
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 553)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY Unpublished (1999)
CONTACT: Walbot V
JOURNAL Department of Biological Sciences
COMMENT Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
Location/Qualifiers
1..553
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"

/clone_lib="660 - Mixed stages of anther and pollen"
/issue_type="whole premiotic anthers to pollen shed"
/dev_stage="premiotic anthers to pollen shed"
/lab_host="X10LR"
/note="Organ: anthers; Vector: Lambda Zap; Site: 1; EcoRI;
site: 2; XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."

BASE COUNT 171 a 111 c 129 g 142 t

Query Match 35.4%; Score 517.8; DB 9; Length 553;
Best Local Similarity 98.7%; Pred. No. 7.6e-66;
Matches 522; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 903 GGACGCGCTTACCAAGAGCTCGACGCTTGTTCAGAGAGCGCTTAATGTCACATTGGATAT 962
|||||
Db 4 GGACGCGCTTACCAAGAGCTCGACGCTTGTTCAGAGAGCGCTTAATGTCACATTGGATAT 63
|||||

QY 963 TCCTGAGCTAAATGAGCTGCACCTGATGAGAGGGTCTCATTAAGTTTCTGTAAAGA 1022
|||||
Db 64 TCCTGAGCTAAATGAGCTGCACCTGATGAGAGGGTCTCATTAAGTTTCTGTAAAGA 123
|||||

QY 1023 TATGTTTCAACGAAGATCGGGTGACAAAGCCATAGAGAAATCAATCTGCCAAGAA 1082
|||||
Db 124 TATGTTTCAATGAAGATCGGGTGACAAAGCCATAGAGAAATCAATCTGCCAAGAA 183
|||||

QY 1083 TAAATGCTGCAAGAGAGACTGAGTCTTTTTCAGGCAACTGCCACATCAGCACC 1142
|||||
Db 184 TAAATGCTGCAAGAGAGACTGAGTCTTTTTCAGGCAACTGCCACATCAGCACC 243
|||||

QY 1143 GCTAAACGGAAGAGACTTCGATTAATAACAAAGCAGCAGCTCGAAGAAACAA 1202
|||||
Db 244 GCTAAACGGAAGAGACTTCGATTAATAACAAAGCAGCAGCTCGAAGAAACAA 303
|||||

QY 1203 GGCTGTGGAAGAGAAATATATCTTGATGCTTGATGATACACTAGCAGTACGAAGA 1262
|||||
Db 304 GGCTGTGGAAGAGAAATATATCTTGATGCTTGATGATGATACACTAGCAGTACGAAGA 363
|||||

QY 1263 GCGGTGCGGTGATCACTTCGTTAGATTATTTAACTCCCTTTTAACTCGAGAGCTTGG 1322
|||||
Db 364 GCGGTGCGGTGATCACTTCGTTAGATTATTTAACTCCCTTTTAACTCGAGAGCTTGG 423
|||||

QY 1323 TAAAGTTTGCATGTTTCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATTGGTGA 1382
|||||
Db 424 TGAAGTTTGCATGTTTCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATTGGTGA 483
|||||

QY 1383 CCAGTAACAAACTATGCGTTTAACTCTGTGCTTTGAAGA 1431
|||||
Db 484 CCAGTAACAAACTATGCGTTTAACTCTGTGCTTTGAAGA 532
|||||

RESULT 6
AM000375 532 bp mRNA linear EST 08-SEP-1999
LOCUS 614014D03.y1.614 - root cDNA library from Walbot Lab zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AM000375
VERSION AM000375.1 GI:5847296
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 532)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY Unpublished (1999)
CONTACT: Walbot V
JOURNAL Department of Biological Sciences
COMMENT Stanford University


```

RESULT 8
AM559173/c 474 bp mRNA linear EST 07-MAR-2000
LOCUS 660065H06.xl 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AM559173
VERSION AM559173
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 474)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
Location/Qualifiers
1..474
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 113 a 114 c 92 g 154 t 1 others
ORIGIN
Query Match 31.7%; Score 463.4; DB 9; Length 474;
Best Local Similarity 98.5%; Pred. No. 5.9e-58;
Matches 467; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 920 AAGTCGAGCCTGTTCAGAGAGCCTAATGTCACATTGGATATTCCTGACCTAAATGGA 979
|||||
DB 474 AAGTCGAGCCTGTTCAGAGAGCCTAATGTCACATTGGATATTCCTGACCTAAATGGA 415

QY 980 CTCGACCTGATGAGAGGGCTCTCATAGTTTCTCGTGAATAAAGTAATGTTTTCACGAG 1039
|||||
DB 414 CTCGACCTGATGAGAGGGCTCTCATAGTTTCTCGTGAATAAAGTAATGTTTTCACGAG 355

QY 1040 ATCGGCTGACAAAGGCCATAGAGAGATCAATTCGCCAAGATTAATTCGCGAAGAA 1099
|||||
DB 354 ATCGGCTGACAAAGGCCATAGAGAGATCAATTCGCCAAGATTAATTCGCGAAGAA 295

QY 1100 GACTCGAGTCCTTTTCAAGCCCACTGCCACATCAGCAGCGCTTAAACGAGAGAGA 1159
|||||
DB 294 GACTCGAGTCCTTTTCAAGCCCACTGCCACATCAGCAGCGCTTAAACGAGAGAGA 235

QY 1160 CTTGCGATAAACAAGCAGCAGCTGCGAACAAAGAAAGGCTGGTGAAGAAGA 1219
|||||
DB 234 CTTGCGATAAACAAGCAGCAGCTGCGAACAAAGAAAGGCTGGTGAAGAAGA 175

QY 1220 AATATATTTGGATGCTTGATGTACACTACGACTACGAAGACAGCGTGGCGATGACT 1279
|||||
DB 174 AATATATTTGGATGCTTGATGTACACTACGACTACGAAGACAGCGTGGCGATGACT 115

QY 1280 TCGCTAGATTATTAACCTCCCTGTTTAACTCAGAGCCTTGTGTAAGATTGCTCATGT 1339

```

```

|||||
DB 114 TCGCTAGATTATTAACCTCCCTGTTTAACTCAGAGCCTTGGTGAAGTTGGCCAGT 55
QY 1340 TTCAACCTGGGTAAGTACTGTTGTTGAAGAGATTGGTACCAAGTACAA 1393
|||||
DB 54 TTCAACCTGGGTAAGTACTGTTGTTGAAGAGATTGGTACCAAGTACAA 1

```

```

RESULT 9
A1861468/c 470 bp mRNA linear EST 19-JUL-1999
LOCUS 614014D03.xl 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION A1861468
VERSION A1861468.1 GI:5525575
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 470)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614014 row: D column: 03.
Location/Qualifiers
1..470
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOLR"
/note="Organ: root; Vector: pBluescriptII SK+; Site_1:
EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
BASE COUNT 112 a 113 c 93 g 152 t
ORIGIN
Query Match 31.6%; Score 462; DB 9; Length 470;
Best Local Similarity 98.9%; Pred. No. 9.5e-58;
Matches 465; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 937 AAGGACCTAATGTGCATGATGATTTCTGAGCTAAATGACCTGCCTGATGAGAG 996
|||||
DB 470 AAGGACCTAATGTGCATGATGATTTCTGAGCTAAATGACCTGCCTGATGAGAG 411

QY 997 GGTCTCATAGTTTCTCGTGAATAAAGATAGTTTCAACGAGATCGGGTACAAAGGCC 1056
|||||
DB 410 GGTCTCATAGTTTCTCGTGAATAAAGATAGTTTCAACGAGATCGGGTACAAAGGCC 351

QY 1057 ATAGAGAGATCAATCTGCCAAGATTAATGTCGCGAAGAAAGACTCGAGTCTTTTTC 1116
|||||
DB 350 ATAGAGAGATCAATCTGCCAAGATTAATGTCGCGAAGAAAGACTCGAGTCTTTTTC 291

QY 1117 AAGCCAAGTCCACCATCATCAGCACCGCTAAACGGAAGAGACTTCGATAAACAAC 1176
|||||
DB 290 AAGCCAAGTCCACCATCATCAGCACCGCTAAACGGAAGAGACTTCGATAAACAAC 231

QY 1177 AAGGCACTGCGAACAAGAAACAAGGCTGGTGAAGAAGAAATATCTTGGATGCTT 1236
|||||
DB 230 AAGGCACTGCGAACAAGAAACAAGGCTGGTGAAGAAGAAATATCTTGGATGCTT 171

```

QY 1237 GATGTACACTAGACAGAGAGGGGTGGCTGATCTGCTTAGATTATTTAA 1296
 |||||||
 Db 170 GATGTACACTAGACAGAGAGGGGTGGCTGATCTGCTTAGATTATTTAA 111
 |||||||
 QY 1297 CTCCCTGTTTACTCAGAGCTTTGTTAAAGTTTGTCTCATGTTTCAAGCTGGGGTAAGT 1356
 |||||||
 Db 110 CTCCCTGTTTAACTCAGAGCTTTGTTAAAGTTTGTCTCATGTTTCAAGCTGGGGTAAGT 51
 |||||||
 QY 1357 TACTGTGTTTAAAGAGATTGTTGACCAAGTAACTATTCGCTGT 1406
 |||||||
 Db 50 TACTGTGTTTAAAGAGATTGTTGACCAAGTAACTATTCGCTGT 1

RESULT 10
 BE186786/c 467 bp mRNA linear EST 22-JUN-2000
 LOCUS 946012C08.XI.946 - tassal primordialium prepared by schmidt lab Zea
 DEFINITION mays cDNA, mRNA sequence.
 ACCESSION BE186786
 VERSION BE186786.1 GI:8665970
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 467)
 Walbot, V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 CONTACT: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 723 8221
 Email: walbot@stanford.edu
 Plate: 946012 row: C column: 08.
 Location/Qualifiers
 1. 467

FEATURES
 source
 /organism="Zea mays"
 /cultivar="OH43"
 /db_xref="taxon:4577"
 /clone_lib="946 - tassal primordialium prepared by Schmidt
 lab"
 /issue_type="tassels"
 /dev_stage="just after the transition from vegetative to
 inflorescence development"
 /lab_host="XLOLR"
 /note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
 Site_2: XhoI; George Chuck dissected immature tassels
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA
 library in HybriZAP. Sample insert size range was 350 bp
 to 3 Kb with a 1 Kb average."
 BASE COUNT 108 a 110 c 96 g 153 t
 ORIGIN

Query Match 30.6%; Score 448; DB 9; Length 467;
 Best Local Similarity 98.7%; Pred No. 1e-55; Mismatches 5; Indels 1; Gaps 1;
 Matches 462; Conservative 0;

QY 886 AGATATCAATTCCTGAGAGCTGGCTTACCAAGAGCTCGAGCTTGTTCAGAGAGCCT 945
 |||||||
 Db 467 AGATATCAATTCCTGAGAGCTGGCTTACCAAGAGCTCGAGCTTGTTCAGAGAGCCT 408
 |||||||
 QY 946 AATGTCAATTCCTGAGAGCTGGCTTACCAAGAGCTCGAGCTTGTTCAGAGAGCCT 1005
 |||||||
 Db 407 AATGTCAATTCCTGAGAGCTGGCTTACCAAGAGCTCGAGCTTGTTCAGAGAGCCT 348
 |||||||
 QY 1006 AGTTTCTGTAAAGATTAAGTTTCAACAGAGATCGGTGACAAAGCCATAGAGAG 1065
 |||||||
 Db 347 AGTTTCTGTAAAGATTAAGTTTCAACAGAGATCGGTGACAAAGCCATAGAGAG 288
 |||||||

QY 1066 ATCAATTCGCAAGATTAATGCTGCAAGAGAGCTGAGTCTTTTCAAGCAGCACT 1125
 |||||||
 Db 287 ATCAATTCGCAAGATTAATGCTGCAAGAGAGCTGAGTCTTTTCAAGCAGCACT 228
 |||||||
 QY 1126 GCCACACATGACGACCGCTTAAAGAGAGAGAGCTGAGTAAACAGAGAGCAGCT 1185
 |||||||
 Db 227 GCCACACATGACGACCGCTTAAAGAGAGAGAGCTGAGTAAACAGAGAGCAGCT 168
 |||||||
 QY 1186 GCGAACAGAAAGAAAGAGCTGCTGGGAAAGAGAGATTAATCTTGTGATGCTTATGTACAA 1245
 |||||||
 Db 167 GCGAACAGAAAGAAAGAGCTGCTGGGAAAGAGAGATTAATCTTGTGATGCTTATGTACAA 108
 |||||||
 QY 1246 CTACGACTACGAAAGAGAGCGGTGCGTATCTGCTTAGATTATTAATCTGCTGT 1305
 |||||||
 Db 107 CTACGACTACGAAAGAGAGCGGTGCGTATCTGCTTAGATTATTAATCTGCTGT 49
 |||||||
 QY 1306 TTAATCTGAGAGCTTTGGTAAAGATTGCTCATGCTTTTCAGCTGGGGTA 1353
 |||||||
 Db 48 TTAATCTGAGAGCTTTGGTAAAGATTGCTCATGCTTTTCAGCTGGGGTA 1

RESULT 11
 AWS62517/c 475 bp mRNA linear EST 10-MAR-2000
 LOCUS 660065H06.X2.660 - Mixed stages of anther and pollen Zea mays cDNA,
 DEFINITION mRNA sequence.
 ACCESSION AWS62517
 VERSION AWS62517.1 GI:7216395
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 475)
 Walbot, V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 CONTACT: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 723 8221
 Email: walbot@stanford.edu
 Plate: 660065 row: H column: 06.
 Location/Qualifiers
 1. 475

FEATURES
 source
 /organism="Zea mays"
 /cultivar="Oh1043"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /issue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="XLOLR"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Anne Franklin.
 BASE COUNT 124 a 110 c 94 g 147 t
 ORIGIN

Query Match 29.4%; Score 429.8; DB 9; Length 475;
 Best Local Similarity 98.2%; Pred. No. 4.3e-53; Mismatches 7; Indels 1; Gaps 1;
 Matches 445; Conservative 0;

QY 979 ACTGACCTGATGAGAGAGGCTGCTAATGTTCTGTTAAAGATTAATGTTTCAACGAA 1038
 |||||||
 Db 475 ACTGACCTGATGAGAGAGGCTGCTAATGTTCTGTTAAAGATTAATGTTTCAACGAA 416
 |||||||
 QY 1039 GATCGGTGACCAAGAGAGAGATTAATCTGCCAAGATTAATGTTGCAAGGA 1098
 |||||||

```

Db      415  |||||||
          GATCGGGGTACAAAGCCATAGAGATCAAAATCGCAAGATTAATCGTCCCAAGGA 356
QY      1099  AGACGTGATCCCTTTTCAAGCCCACTCCACACATCAGCAGCGTAAACGGAAGAG 1158
          |||||||
Db      355  AGACTGATCCCTTTTCAAGCCCACTCCACACATCAGCAGCGTAAACGGAAGAG 296
QY      1159  ACTTCGGATTAACCAAGCAAGCAGCTGCGAACAAGAAACAAAGCGTGTGGAAGAG 1218
          |||||||
Db      295  ACTTCGGATTAACCAAGCAAGCAGCTGCGAACAAGAAACAAAGCGTGTGGAAGAG 236
QY      1219  AATATATCTTGATGCTTGATGATACACTACAGCTACAGCAAGCGGTGCGATGATC 1278
          |||||||
Db      235  AATATATCTTGATGCTTGATGATACACTACAGCTACAGCAAGCGGTGCGATGATC 176
QY      1279  TTGCGTAAATATTAATCTCCCTGTTTAACTCAGACTTGTGTAAGAGTTCATG 1338
          |||||||
Db      175  TTGCGTAAATATTAATCTCCCTGTTTAACTCAGACTTGTGTAAGAGTTCATG 116
QY      1339  TTTCAAGCTGGGGTAACTAGTGTGTTGAAGAGATTGTACCAAGTAAACAACTT 1398
          |||||||
Db      115  TTTCAAGCTGGGGTAACTAGTGTGTTGAAGAGATTGTACCAAGTAAACAACTT 57
QY      1399  ATGCGTGTGTTTACTCTCTCTCTCTTTGAAGTA 1431
          |||||||
Db      56  ATGCGTGTGTTTACTCTCTCTCTTTGAAGTA 24

```

```

RESULT 12
AM774700 733 bp mRNA linear EST 07-SEP-2000
LOCUS     AM774700
DEFINITION
sequence.
ACCESSION AM774700
VERSION   AM774700.1 GI:7718617
KEYWORDS  EST.
SOURCE    barrel medic.
ORGANISM  Medicago truncatula
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
           Medicago.
REFERENCE 1 (bases 1 to 733)
           Vandenbosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L.,
           Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
           Fraser,C.M.
           ESTs from roots of Medicago truncatula after Rhizobium inoculation
           Unpublished (1999)
           Contact: Vandenbosch K
           Department of Biology
           Texas A&M University
           College Station, TX 77843-3258, USA
           Tel: 409 845 7707
           Fax: 409 845 2891
           Email: kate@mail.bio.tamu.edu
           Texas A&M EST name: T258119e
           TIGR sequence name: MTEBE71TK
           More information is available at:
           http://chryslie.tamu.edu/medicago
           Seq primer: Skmod (CTA GAA CTC GAT CC).
           Location/Qualifiers
             1..733

```

```

FEATURES
  source
    /organism="Medicago truncatula"
    /cultivar="genotype A17"
    /db_xref="taxon:3880"
    /clone="PKV3-23L21"
    /clone_11b="KV3"
    /tissue_type="Seedling roots"
    /dev_stage="3 days post-inoculation with sinorhizobium
    meliloti"
    /lab_host="E. coli strain XLOLR"
    /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
    XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA

```

```

was directionally ligated into the Unizap XR vector from
Stratagene and packaged using GigaPack IIT Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XLOLR cells."
BASE COUNT      222 a      125 c      183 g      203 t
ORIGIN
Query Match      28.7%; Score 420.4; DB 9; Length 733;
Best Local Similarity 75.4%; Pred. No. 8.3e-52;
Matches 523; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

```

```

84 GATGGGATCAAGGGTTGACCAACTGTCGCGCAATATGCCGCCAAGCGATGAAGA 143
Db      40  GATGGGATTAAGGGTTTAACGAAGCTTTTACGTGATATGCTCCCAATGAAGA 99
QY      144  GCAGAAGTTCCGAGACTCTTCGGCCGCAAAATGCCGTGAGCGCCAGCTAGCATATA 203
          |||||||
Db      100  GAACAATTCGAATCTTACTTGGCGTAAGATTGCTGTGATGCTAGTGAAGCATTTA 159
QY      204  CCAGTCTCGATTGATTGATTGACAGCAGCATGGAACCTCACAATGAAAGCTGTGA 263
          |||||||
Db      160  CCAGTCTCTATTGTTGTTGGGAGAGAGTGAAGTGAATGTTGATGATGAAGCTGTGA 219
QY      264  AATCACTAGTCAATTTGCAAGCAATGTTCAACCGACATTAAGATTACTGAGACGGGAAT 323
          |||||||
Db      220  ACTAATAGTCAATTTGCAAGCAATGTTGCGCGGACATCAAGCTCTTCAAGACCGGGAT 279
QY      324  CAAGCAGTTATGTTTGTGATGCGACCTCCTGATATGAGCAAGCAAGAGCTTCTTA 383
          |||||||
Db      280  GAAGCAGATATATGTTTGTGATGGAAGCCACCGGATGAAGATCAAGAGCTGAAGAA 339
QY      384  AAGATATCAAAAAGATGATGACACCAAGATCTGACGTGAGGCACTAAGAGTAGAGAG 443
          |||||||
Db      340  ACGTCTCTCAAGAGAGCTGAGGCTACCGCGGTTTACAGAACTCTAGAGGCTGACA 399
QY      444  TAAAGATGCGATTGAAATTTAGCAGAGAGGACTGTAAAGTCTCAAGCAAGCAAGCA 503
          |||||||
Db      400  TAAGAGATGATTAAGAAATTTACATTAACGACAGTGAAGTGAACCAATATATGA 459
QY      504  AATTTGTAACGGCTATTAAGCTTATGAGGGTCTCGTGTGATGAGCAGCTCTGTAGC 563
          |||||||
Db      460  CCAGTCTCAAAAGACTTTTGAGACATGAGAGTGCCTGTTGTGAGCAGCTCAGAGAG 519
QY      564  AAGACAGAGATGCAAGCCCTTTGCAATAAAGATAGAGTGTGCTGTCTTCAGAGAG 623
          |||||||
Db      520  AGAGGCTAGTGTGCTGCGACTTGGCAAGCTGGAAGGTGATGCTGTGCTCAAGA 579
QY      624  TATGACATCCCTTACTTCTTGGGGCTCCAGGCTTCCTGCTGATTTAATGATGAGAG 683
          |||||||
Db      580  CATGATTCCTTAACGTTTGGAGCTCCTAAGTTCCTTCGCGCATTAATGATGATCTAG 639
QY      684  CAGAAATACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 743
          |||||||
Db      640  AAGAGAGATTCAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 699
QY      744  CAGGACAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
          |||||||
Db      700  CCGGACCAATTTATGACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 733

```

```

RESULT 13
AI834484 553 bp mRNA linear EST 02-FEB-2000
LOCUS     AI834484
DEFINITION
mays cDNA, mRNA sequence.
ACCESSION AI834484
VERSION   AI834484.1 GI:5468693
KEYWORDS  EST.
SOURCE    Zea mays.
ORGANISM  Zea mays.
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

```

clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 553)
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 CONTACT: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 8221
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 606068 row: G column: 09.
 Location/Qualifiers
 1..553
 /organism="Zea mays"
 /cultivar="Oh1043"
 /db_xref="taxon:4577"
 /clone_lib="606 - Ear tissue cDNA library from Schmidt
 lab"
 /issue_type="mixed"
 /dev_stage="ear length from 0.5 cm - 2.0 cm"
 /lab_host="X10LR (Stratagene)"
 /note="Organ: immature ear; Vector: PBK-CMV; Site_1: EcoRI
 ; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
 lab"
 1ab 112 c 106 g 195 t

BASE COUNT 140 a 112 c 106 g 195 t

ORIGIN

Query Match 28.2%; Score 413; DB 9; Length 553;
 Best Local Similarity 96.6%; Pred. No. 1.1e-50;
 Matches 422; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 995 AGGAGCTATAGTTCTCTGTGTAAGATATGCTCAACGAAGATCGGGGCAAAAG 1054
 |||||||
 DB 553 AGGCTCTATAGTTCTCTGTGTAAGATATGCTTCAATGAAGATCGGGGCAAAAG 494

QY 1055 CCATGAGAGAGTCAATCTGCCAAGATTAATCGTGCAGAGAACTCGAGTCTTTT 1114
 |||||||
 DB 493 CCATGAGAGAGTCAATCTGCCAAGATTAATCGTGCAGAGAACTCGAGTCTTTT 434

QY 1115 TCAAGCCAACTGCGACACATCAGACCGCTTAAGAGGAAGAGACTTCGATAAACA 1174
 |||||||
 DB 433 TCAAGCCAACTGCGACCGCCCATCAGCCCGCTTAAGAGGAAGAGCTTCGATAAACA 374

QY 1175 GCAAGGCGAGTGGCAACAAAGAAAGAGCTTGAAGAAAGAAATATCTTGATGC 1234
 |||||||
 DB 373 GCAAGGCGAGTGGCAACAAAGAGCTTGAAGAAAGAAATATCTTGATGC 314

QY 1235 TTGATGTACAACTACGACTACGAAAGCAGCGGTGATCATCTTCCTGATTTT 1294
 |||||||
 DB 313 TTGATGTACAACTACGACTACGAAAGCAGCGGTGATCATCTTCCTGATTTT 254

QY 1295 AACTGCCGTGTTTAACTCAGAGCTTTGTTAAAGTTTGCATGTTTCAAGTGGGGTAA 1354
 |||||||
 DB 253 AACTGCCGTGTTTAACTCAGAGCTTTGTTAAAGTTTGCATGTTTCAAGTGGGGTAA 194

QY 1355 GTTACTGTGTGTTGAAGAGATGCTGACCAAGTAAACAACTTATCGCTGTTTACT 1414
 |||||||
 DB 193 GTTACTGTGTGTTGAAGAGATGCTGACCAAGTAAACAACTTATCGCTGTTTACT 134

QY 1415 TCTTGCTCTTGAAGTA 1431
 |||||||
 DB 133 TCTTGCTCTTGAAGTA 117

RESULT 14
 BG263233 539 bp mRNA linear EST 16-FEB-2001
 LOCUS BG263233
 DEFINITION WHE2339_F04.L07Z5 Wheat pre-anthesis spike cDNA library Trilicium
 aestivum cDNA clone WHE2339_F04.L07, mRNA sequence.
 ACCESSION BG263233

VERSION BG263233.1 GI:12865185
 KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Trilicium aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triliceae; Trilicium.
 REFERENCE 1 (bases 1 to 539)
 Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
 ,P.S., Hsiao,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
 Seaton,C.L. and Tong,J.C.
 The structure and function of the expressed portion of the wheat
 genomes - Pre-anthesis spike cDNA library
 Unpublished (2000)
 JOURNAL Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oanderson@w.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seq primer: Stragene, SK primer.
 FEATURES
 source
 1..539
 Location/Qualifiers
 /organism="Trilicium aestivum"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone_lib="WHE2339_F04.L07"
 /clone_lib="Wheat pre-anthesis spike cDNA library"
 /issue_type="Spike before anthesis"
 /dev_stage="Adult plant"
 /lab_host="E. coli SOLR"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
 greenhouse. Whole spike with awns trimmed, white, green
 and yellow anther were collected and total RNA, and
 poly(A) RNA were prepared, a cDNA library was made, and
 the cDNA clones were in vivo excised to give pluscript
 phagemids in the T3 Close lab (Choi, Close, Fenton) at
 the University of California, Riverside. Plasmid DNA
 preparations and DNA sequencing were performed in the OD
 Anderson lab (all other authors)."
 BASE COUNT 163 a 108 c 141 g 126 t 1 others

ORIGIN

Query Match 28.1%; Score 411.6; DB 10; Length 539;
 Best Local Similarity 85.2%; Pred. No. 1.8e-50;
 Matches 459; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 160 TACTTCGCGCGCAAAATCGCCGTGACGCGACATGATATACAGTCTCGATTGTA 219
 |||||||
 DB 1 TACTTCGCGCGCGCATCGCCGTGACGCGACATGATATACAGTCTCGATTGTA 60

QY 220 GTTGAAGGACAGGCGATGGAACCTCTCAAAATGAAGTGTGAAGTCACTAGTCAATTG 279
 |||||||
 DB 61 GTTGAAGGACAGGCGATGGAACCTCTTACAAAGCAAGCGGTGATACACATTTG 120

QY 280 CAAGCAATGTTCAACCGCATATAGATTCTGGAAGCGGATCAAGCAAGTTATGTT 339
 |||||||
 DB 121 CAAGCAATGTTCAACCGCATATAGATTCTGGAAGCGGATCAAGCAAGTTATGTT 180

QY 340 TTTGATGCGAAGCTCTCTGTATGAGAAACAAGAGCTTGTAAATATCTCAAAAAGA 399
 |||||||
 DB 181 TTTGATGCGAAGCTCTCTGTAAATGAAAGAGCAGCTTTTAAACACACGCAAGAG 240

QY 400 CATGATCAACCAAGATCTGACTGAGCGAGTACAGAGTACAGATTAAGTCCGATTGAA 459
 |||||||
 DB 241 AATTAACCAACAGAGCTCTGAGGAGAGTACAGAGTACAGATTAAGTCCGATTGAA 300

QY 460 AATTTGACAGAGGAGCTGTAAGGTACACAAGCAACGCAAGAGTTTAAACGCGCTA 519
 |||||||

```

Db 301 AATTCAGCAGAGAACTTAAGCTACAGACAGACACATGATGTTTANGCCTCA 360
QY 520 TTAAAGCTATGGGGTTCCTGTTTGTAGAGCACCCTTGTAGACAGACAGATGTGA 579
Db 361 CTAGACGTGTGGTCTGCTGTTGTGAGGCTCCCTGTGAGCAATATCATATGTGCT 420
QY 580 GCCCTTTCATTAAGAGTGTTCCTGTTGCTTCAGAAAGATATGAGCTCCCTTACT 639
Db 421 GCCCTTTCAGAAATGACAAAGGTATGCTGTGATCCCAAGATATGAGCTCCTTACT 480
QY 640 TTTGGGGTCCACGTTCTCTGCTATTAATGATCCCAAGTTCAGAAATACCTGT 698
Db 481 TTTGGAGCTACACGTTTGTGCTATTTGATGATTCACAGGTTCAGAAATACCTGT 539

RESULT 15
AM288784 414 bp mRNA linear EST 16-JAN-2000
LOCUS 707010C02.x5.707 - Mixed adult tissues from Walbot lab (SK) Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION AM288784.1 GI:6695706
VERSION AM288784.1
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays;
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 414)
REFERENCE Walbot, Y.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
AUTHORS University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: C column: 02.
location/Qualifiers
source
1.414
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
}
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 120 a 82 c 101 g 109 t 2 others
ORIGIN

Query Match 27.5%: Score 402.2; DB 9; Length 414;
Best Local Similarity 98.8%: Pred No 4.4e-49;
Matches 404; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 649 CCACGGTTCCTGCTCATTTAATGATCCCAAGTTCAGAAATACCTGTGATGAAATTT 708
Db 6 CCACGGTTCCTGCTCATTTAATGATCCCAAGTTCAGAAATACCTGTGATGAAATTT 65
QY 709 GATGTTCGAAGGTTTGGAGAGCTTGAAGTCAACATGACCAATCATTTGATTTGTC 768
Db 66 GATGTTCGAAGGTTTGGAGAGCTTGAAGTCAACATGACCAATCATTTGATTTGTC 125
QY 769 ATCCTGTGTGATGTGACTATTGTGATACATCAAGATATCGGGGCAAAACAGCTCTG 828

```

```

Db 126 ATCCTGTGTGATGTGACTATTGTGATAGCATCAAGGATATCGGGGCAAAACAGCTCTG 185
QY 829 AAACCTATTGCTCAACATGGGTCATAGAAAGCATCTGAGAAATCTTAATTAAGACAGA 888
Db 186 AAACCTATTGCTCAACATGGGTCATAGAAAGCATCTGAGAAATCTTAATTAAGACAGA 245
QY 889 TATCAAAATTCCTGAGAGCTGACCTTACCAGAAAGCTGACGCTTTGTCAGAGCCTAAT 948
Db 246 TACCAAAATTCCTGAGAGCTGACCTTACCAGAAAGCTGACGCTTTGTCAGAGCCTAAT 305
QY 949 GTACATTTGATATTCCTGAGCTTAAATGACATGACCTGATGAGAGGCTCTCAATAAGT 1008
Db 306 GTACATTTGATATTCCTGAGCTTAAATGACATGACCTGATGAGAGGCTCTCAATAAGT 365
QY 1009 TTCTGTGTAAGATTAATGCTTTTCACAGAGATCGGTGACAAAGGCCA 1057
Db 366 TTCTGTGTAAGATTAATGCTTTTCACAGAGATCGGTGACAAAGGCCA 414

```

Search completed: November 5, 2002, 13:36:44
Job time : 1271.78 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:23:34 (Search time 31 Seconds
(without alignments)
1357.967 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939
Sequence: 1 MGIGLRLKLLADNAPKAKE.....SDRTSKAAANKTKAGKKK 379

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	379	21	AAV95307
2	1939	100.0	379	21	AAV95309
3	1939	99.7	379	21	AAV95308
4	1939	99.7	379	21	AAV95310
5	1031	53.2	380	20	AAW92504
6	1007.5	52.0	377	20	AAW92505
7	998.5	51.5	385	22	ABB63960
8	996	51.4	378	20	AAW92508
9	938	48.4	382	20	AAW92506
10	699.5	36.1	373	22	ABG19545
11	648.5	33.4	340	18	AAW24216

12	648.5	33.4	340	19	AAW79970	Pyrococcus furiosu
13	648.5	33.4	340	19	AAW59940	Amino acid sequenc
14	647	33.4	326	19	AAW59953	Amino acid sequenc
15	644	33.2	332	19	AAW59951	Amino acid sequenc
16	636	32.8	342	22	AAW66644	Putative P. abyssi
17	632.5	32.6	343	22	AAW03778	Pyrococcus heat re
18	609	31.4	325	19	AAW59950	Amino acid sequenc
19	600	30.9	340	19	AAW59949	Amino acid sequenc
20	562	29.0	340	19	AAW59952	Amino acid sequenc
21	560.5	28.9	326	18	AAW24215	Methanococcus jann
22	560.5	28.9	326	19	AAW59939	Amino acid sequenc
23	533.5	26.5	326	19	AAW79969	Methanococcus jann
24	538.5	27.8	336	19	AAW79982	Archaeoglobus fulg
25	538.5	27.8	336	19	AAW59946	Amino acid sequenc
26	525	27.1	328	19	AAW59948	Amino acid sequenc
27	411	21.2	258	19	AAW59947	Amino acid sequenc
28	340.5	17.6	386	20	AAW92507	Yeast delta-RAD2 p
29	336	17.3	1516	21	AAW18195	Plasmodium falcipla
30	257	13.3	726	22	ABB65192	Drosophila melanog
31	243.5	12.6	1236	22	ABB65287	Drosophila melanog
32	243.5	12.6	1257	22	ABB67273	Drosophila melanog
33	210	10.8	732	22	ABB71917	Thermococcus litor
34	208	10.7	872	20	AAV31815	Thermococcus litor
35	205	10.6	872	18	AAW26605	Thermococcus litor
36	205	10.6	872	20	AAV31812	Thermococcus litor
37	205	10.6	872	20	AAV31816	Thermococcus litor
38	184.5	9.5	291	20	AAW78466	T. calidophilus DNA
39	181	9.3	832	13	AAW22602	Taq polymerase enc
40	174.5	9.0	832	13	AAW22603	Taq polymerase enc
41	174.5	9.0	832	13	AAW22604	Taq polymerase enc
42	173.5	8.9	834	17	AAW9542	Thermus flavus DNA
43	173.5	8.9	834	17	AAW96203	DNA polymerase I h
44	173	8.9	548	17	AAW96263	Mutant Thermus aqu
45	173	8.9	695	17	AAW96264	Mutant Thermus aqu

ALIGNMENTS

RESULT 1
ID AAV95307 standard; Protein: 379 AA.
AC AAV95307;
DT 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 protein.
XX
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting.
XX
OS Zea mays.
XX
PN WO200036109-A1.
XX
PD 22-JUN-2000.
XX
PF 16-NOV-1999; 99WO-US27147.
XX
PR 15-DEC-1998; 98US-0112332.
XX
PA (PION) PIONEER HI-BRED INT INC.
XX
PI Mahajan PB;
XX
DR WPI: 2000-452026/39.
XX
DR N-PSDB: AAA27923.
XX
PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX

PS Claim 11; Page 71-72; 85bp; English.
XX
XX The present sequence is that of maize Rad2/FEN-1, as deduced from
CC a cDNA clone (see AAA27923) isolated from maize line B73 immature
CC ear tissue. Rad2/FEN-1 is a structure specific endonuclease which
CC can be expressed in transgenic plant cells using conventional
CC methods. The protein is involved in the regulation of DNA repair
CC and recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 379 AA:

Query Match 100.0%; Score 1939; DB 21; Length 379;
Best local Similarity 100.0%; Pred. No. 1.2e-170;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKIGLTKLLADNAPKAMEKQKFESYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
DB 1 MGKIGLTKLLADNAPKAMEKQKFESYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
QY 61 VTSHLQGMFNRTIRLLEAGIKPYVFDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120
DB 61 VTSHLQGMFNRTIRLLEAGIKPYVFDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120
QY 121 KDAIEKLSKRTYAVTROHNECDCKRLRLMGVPVPAFSAEACALCINIKVAVASD 180
DB 121 KDAIEKLSKRTYAVTROHNECDCKRLRLMGVPVPAFSAEACALCINIKVAVASD 180
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMOQFIDLCILCGCYCDSIK 240
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMOQFIDLCILCGCYCDSIK 240
QY 241 GIGGQTAALKLIRQHGSIESTLENLNDKRYQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300
DB 241 GIGGQTAALKLIRQHGSIESTLENLNDKRYQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300
QY 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
DB 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAANKTKYAGGKK 379
DB 361 DKTSKAANKTKYAGGKK 379

RESULT 2
AAAY95309 standard; Protein; 379 AA.
XX
XX AAAY95309;
XX 12-SEP-2000 (first entry)
XX
XX Maize Rad2/FEN-1 protein.
XX
XX Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
XX endonuclease; exonuclease; DNA repair; gene targeting.
XX
XX Zea mays.
XX
XX WO200036109-A1.
XX
XX 22-JUN-2000.
PD

XX
XX 16-NOV-1999; 99WO-US27147.
XX
XX 15-DEC-1998; 98US-0112332.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Mahajan PB;
XX
XX WPI: 2000-452026/39.
XX N-PSDB: AAA27925.
XX
XX Example 1; Page 76-77; 85bp; English.
XX
XX The present sequence is that of maize Rad2/FEN-1, as deduced from
CC a cDNA clone (see AAA27925) derived from maize line W23 tassel
CC polyA RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC can be expressed in transgenic plant cells using conventional
CC methods. The protein is involved in the regulation of DNA repair
CC and recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 379 AA:

Query Match 100.0%; Score 1939; DB 21; Length 379;
Best local Similarity 100.0%; Pred. No. 1.2e-170;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKIGLTKLLADNAPKAMEKQKFESYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
DB 1 MGKIGLTKLLADNAPKAMEKQKFESYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
QY 61 VTSHLQGMFNRTIRLLEAGIKPYVFDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120
DB 61 VTSHLQGMFNRTIRLLEAGIKPYVFDGKPPDMKKQELAKRYSKRDATKDLTEAVEYGD 120
QY 121 KDAIEKLSKRTYAVTROHNECDCKRLRLMGVPVPAFSAEACALCINIKVAVASD 180
DB 121 KDAIEKLSKRTYAVTROHNECDCKRLRLMGVPVPAFSAEACALCINIKVAVASD 180
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMOQFIDLCILCGCYCDSIK 240
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDVAVLELELTMOQFIDLCILCGCYCDSIK 240
QY 241 GIGGQTAALKLIRQHGSIESTLENLNDKRYQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300
DB 241 GIGGQTAALKLIRQHGSIESTLENLNDKRYQIPEDMPYQEARLFEKPNVTLDIPELKWTA 300
QY 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
DB 301 PDEEGLISFLVKNNGFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAANKTKYAGGKK 379
DB 361 DKTSKAANKTKYAGGKK 379

RESULT 3
AAAY95308


```

ID  AAY95308 standard; Protein; 379 AA.
XX
AC  AAY95308;
XX
DT  12-SEP-2000 (first entry)
XX
DE  Maize Rad2/FEN-1 protein.
XX
XX  Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KM  endonuclease; exonuclease; DNA repair; gene targeting.
XX
OS  Zea mays.
XX
PN  WO200036109-A1.
XX
PD  22-JUN-2000.
XX
PF  16-NOV-1999; 99WO-US27147.
XX
PR  15-DEC-1998; 98US-0112332.
XX
PA  (PION-1) PIONEER HI-BRED INT INC.
XX
PI  Mahajan PB;
XX
DR  WPI: 2000-452026/39.
XX  N-PSDB; AAA27924.
XX
PS  Example 1; Page 74; 85pp; English.
XX
CC  The present sequence is that of maize Rad2/FEN-1, as deduced from
CC  a cDNA clone (see AAA27924) derived from maize line B73 seedling
CC  tissue RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC  under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC  can be expressed in transgenic plant cells using conventional
CC  methods. The protein is involved in the regulation of DNA repair
CC  and recombination in plant systems and therefore may be used for
CC  improving gene targeting during further recombinant DNA protocols
CC  involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC  in DNA replication and nucleotide excision and repair reactions.
CC  The exolytic activity is involved in double strand break repair and
CC  end joining. The protein is also useful in strand exchange
CC  reactions during homologous recombination. These functions may be
CC  useful in gene targeting and in the production of male sterile
CC  plants. The efficacy of gene targeting can be improved by the
CC  overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC  be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ  Sequence 379 AA:

Query Match          99.7%: Score 1933; DB 21; Length 379;
Best Local Similarity 99.7%: Pred. No. 4.2e-170;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY  1 MGIKLTFLADNAPKAKKEGFESYFGRIADVASKISYQFIIVGRTGMETLNEAGE 60
DB  1 MGIKLTFLADNAPKAKKEGFESYFGRIADVASKISYQFIIVGRTGMETLNEAGE 60
OY  61 VTSHOGMFPNFTIRLENGITRPVVFPGDKPPDKMKOFLARKYRKRDATKDLTEAVEVD 120
DB  61 VTSHOGMFPNFTIRLENGITRPVVFPGDKPPDKMKOFLARKYRKRDATKDLTEAVEVD 120
OY  121 KDAIEKLSKRTVKYTRQHNEDEKRLRLMGVPVVEAPSEAEACALCINDKFAVASD 180
DB  121 KDAIEKLSKRTVKYTRQHNEDEKRLRLMGVPVVEAPSEAEACALCINDKFAVASD 180
OY  181 MDSLTFGAPRFLRLMDPSSKKIPVMEFDVAKVLEELTMDQFIDICLTCGCDYDSIK 240
DB  181 MDSLTFGAPRFLRLMDPSSKKIPVMEFDVAKVLEELTMDQFIDICLTCGCDYDSIK 240

```

```

OY  241 GIGGOTALKLIROHGSIESILENLKNDRYQIPEDWPYQOEARLREKPNVTLDIPELKMTA 300
DB  241 GIGGOTALKLIROHGSIESILENLKNDRYQIPEDWPYQOEARLREKPNVTLDIPELKMTA 300
OY  301 PDEEGLISFLVNDNGNEPDRVTKATEKIKSAKNKSSQGRLESPFKPATTSAPLKRKETS 360
DB  301 PDEEGLISFLVNDNGNEPDRVTKATEKIKSAKNKSSQGRLESPFKPATTSAPLKRKETS 360
OY  361 DKTSKAANKKTRAGGKK 379
DB  361 DKTSKAANKKTRAGGKK 379

RESULT 4
ID  AAY95310 standard; Protein; 379 AA.
XX
AC  AAY95310;
XX
DT  12-SEP-2000 (first entry)
XX
DE  Maize Rad2/FEN-1 protein.
XX
KM  Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KM  endonuclease; exonuclease; DNA repair; gene targeting.
XX
OS  Zea mays.
XX
PN  WO200036109-A1.
XX
PD  22-JUN-2000.
XX
PF  16-NOV-1999; 99WO-US27147.
XX
PR  15-DEC-1998; 98US-0112332.
XX
PA  (PION-1) PIONEER HI-BRED INT INC.
XX
PI  Mahajan PB;
XX
DR  WPI: 2000-452026/39.
XX  N-PSDB; AAA27926.
XX
PS  Example 1; Page 79-80; 85pp; English.
XX
CC  The present sequence is that of maize Rad2/FEN-1, as deduced from
CC  a cDNA clone (see AAA27926) derived from maize line B73 endosperm
CC  RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC  under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC  can be expressed in transgenic plant cells using conventional
CC  methods. The protein is involved in the regulation of DNA repair
CC  and recombination in plant systems and therefore may be used for
CC  improving gene targeting during further recombinant DNA protocols
CC  involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC  in DNA replication and nucleotide excision and repair reactions.
CC  The exolytic activity is involved in double strand break repair and
CC  end joining. The protein is also useful in strand exchange
CC  reactions during homologous recombination. These functions may be
CC  useful in gene targeting and in the production of male sterile
CC  plants. The efficacy of gene targeting can be improved by the
CC  overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC  be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ  Sequence 379 AA:

Query Match          99.7%: Score 1933; DB 21; Length 379;
Best Local Similarity 99.7%: Pred. No. 4.2e-170;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MGIGLTKLADNAPKAMKEOKFESYFGKRIADVDSMSIYQFLIVGRGTMETLNEAGE 60
    |||
DB 1 MGIGLTKLADNAPKAMKEOKFESYFGKRIADVDSMSIYQFLIVGRGTMETLNEAGE 60
QY 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVD 120
    |||
DB 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVD 120
QY 121 KDAIEKLSRIVKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFAYASED 180
    |||
DB 121 KDAIEKLSRIVKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFAYASED 180
QY 181 MSITPGARFLRLHMDSPSKIPYMEEDVAKVLEELTMDQFIDLCICGCDYCDISK 240
    |||
DB 181 MSITPGARFLRLHMDSPSKIPYMEEDVAKVLEELTMDQFIDLCICGCDYCDISK 240
QY 241 GIGGOTALKLIROHGSIESILENLKRDYQIPEDMPYQEARRLFEKEDNVTLDIPELKWTA 300
    |||
DB 241 GIGGOTALKLIROHGSIESILENLKRDYQIPEDMPYQEARRLFEKEDNVTLDIPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEEDRVTKAIEKTSKAKNSOGRLSEFPKPTATTSAPLRKETS 360
    |||
DB 301 PDEGLISFLVKGNGFNEEDRVTKAIEKTSKAKNSOGRLSEFPKPTATTSAPLRKETS 360
QY 361 DKTSAANKTKTAGKKK 379
    |||
DB 361 DKTSAANKTKTAGKKK 379

RESULT 5
AAM92504
ID AAM92504 standard; Protein: 380 AA.
AC AAM92504;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 protein.
XX
KW FEN-1; human; flap endonuclease; detection: diagnosis: carcinogen;
KW neoplasma; antineoplastic agent; cleavage.
XX
OS Homo sapiens.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
PI Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI: 1999-179985/15.
XX
XX N-PSDB; AAX02107.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
XX e.g. recombinant polypeptides
XX
PS Claim 1; Fig 1A; 58pp: English.
XX
CC This sequence represents a human FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel

```

```

CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 380 AA;
Query Match 53.8%; Score 1031; DB 20; Length 380;
Best local Similarity 53.8%; Pred. No. 9.6e-87;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;
QY 1 MGIGLTKLADNAPKAMKEOKFESYFGKRIADVDSMSIYQFLIVGRGTMETLNEAGE 60
    |||
DB 1 MGIGLTKLADNAPKAMKEOKFESYFGKRIADVDSMSIYQFLIVGRGTMETLNEAGE 60
QY 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVD 120
    |||
DB 61 VTSHLOGMNRITRLLENGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVD 120
QY 121 KDAIEKLSRIVKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFAYASED 180
    |||
DB 121 KDAIEKLSRIVKVTROHNEDCKRLRLMGVYVVEAPSEAEACALCINDKVFAYASED 180
QY 181 MSITPGARFLRLHMDSPSKIPYMEEDVAKVLEELTMDQFIDLCICGCDYCDISK 240
    |||
DB 181 MSITPGARFLRLHMDSPSKIPYMEEDVAKVLEELTMDQFIDLCICGCDYCDISK 240
QY 241 GIGGOTALKLIROHGSIESILENLKRDYQIPEDMPYQEARRLFEKEDNVTLDIPELKWTA 300
    |||
DB 241 GIGGOTALKLIROHGSIESILENLKRDYQIPEDMPYQEARRLFEKEDNVTLDIPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEEDRVTKAIEKTSKAKNSOGRLSEFPKPTATTSAPLRKETS 360
    |||
DB 301 PDEGLISFLVKGNGFNEEDRVTKAIEKTSKAKNSOGRLSEFPKPTATTSAPLRKETS 360
QY 359 TSDK--TSKAAANKTKTAGKKK 379
    |||
DB 358 PEPRGSKKKKA--KTGAAGKFK 377

RESULT 6
AAM92505
ID AAM92505 standard; Protein: 377 AA.
AC AAM92505;
XX
DT 23-APR-1999 (first entry)
XX
DE Mouse FEN-1 protein.
XX
KW FEN-1; mouse; flap endonuclease; detection: diagnosis: carcinogen;
KW neoplasia; antineoplastic agent; cleavage.
XX
OS Mus sp.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
PI Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI: 1999-179985/15.
XX
XX N-PSDB; AAX02108.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing

```

PT e.g. recombinant polypeptides

XX PS Claim 1; Fig 2A; 58pp; English.

CC This sequence represents a mouse FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasms, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.

XX Sequence 377 AA;

Query Match 52.0%; Score 1007.5; DB 20; Length 377;
Best Local Similarity 53.4%; Pred. No. 1.4e-84;
Matches 203; Conservative 68; Mismatches 100; Indels 9; Gaps 7;

QY 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMETLTNEAGE 60
DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMETLTNEAGE 59
QY 61 VTSHLQGMENRTIRLEAGIKRPYVFDGKPPMKKQELAKRYSKRDADTKDLTEAVEVGD 120
DB 60 TTS-LMGFRTTR-MENGIRPYVFDGKPPMKKQELAKRYSKRDADTKDLTEAVEVGD 116
QY 121 KDAIEKLSKRTVAVKROHNEDECKRLRLMGVPEVVEASEAECAALCINDKFAVAASED 180
DB 117 MEEVEKTKRLVAVKROHNEDECKRLRLMGVPEVVEASEAECAALCINDKFAVAASED 176
QY 181 MDSLTFGAPRLRHLMDPPSKKIPVMEPDVAKVLEBELTMDQFIDLCILGCDYCDYSIK 240
DB 177 MDCLTFESPLMKHLTASKEKKIPIDIEFHLRYLOELIGNQEPYDLCILGCDYCESIR 236
QY 241 GIGGQALKILROHGSIESILENLKDRYQIPEDMYPQEARLFEKPNVT-LDIPELKWT 299
DB 237 GIGKRAVVDILQHKSTIEELVRRLDPSKYVPENMLHKEAOQLFEPYVPESEVELKWS 296
QY 300 APDEGLISFLVNDGNGFENDRYTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRKET 359
DB 297 EPNEELVKMGCEKQSEERIRSGVKRLSKSGOSTQGRLDLDFKVGYSLS-AKKREP 355
QY 360 SDKTSAANKKTKAGGKK 379
DB 356 EPKGS--AKKAKGTGAGK 372

RESULT 7

ABB63960 ID ABB63960 standard; Protein; 385 AA.

AC ABB63960;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 18672.

XX Drosophila: developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL08063.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 18672; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130311), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB57737-AB472072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 385 AA;

Query Match 51.5%; Score 998.5; DB 22; Length 385;
Best Local Similarity 52.9%; Pred. No. 9.9e-84;
Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps 5;

QY 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMETLTNEAGE 60
DB 1 MGIGKLTLLADNAPKAMKQKFESEYFGKRIADVADSMSTYQFLIVGRTGMETLTNEAGE 59
QY 61 VTSHLQGMENRTIRLEAGIKRPYVFDGKPPMKKQELAKRYSKRDADTKDLTEAVEVGD 120
DB 60 PSHLMGMFRTIRLDNGIRPYVFDGKPPDLKSGELAKRAREERPEAEKALKAADADG 119
QY 121 KDAIEKLSKRTVAVKROHNEDECKRLRLMGVPEVVEASEAECAALCINDKFAVAASED 180
DB 120 DAGIEFNNRLVAVKRLHAKKAKKELLTLMGVYVDACENACCAALYKRGKYATATED 179
QY 181 MDSLTFGAPRLRHLMDPPSKKIPVMEPDVAKVLEBELTMDQFIDLCILGCDYCDYSIK 240
DB 180 MDALTFGSKRLRYLTYSKARKMPYKESYDKLLEGLAINNREFIDLCILGCDYCESIR 239
QY 241 GIGGQALKILROHGSIESILENLKDRYQIPEDMYPQEARLFEKPNVT-LDIPELKWT 299
DB 240 GIGKRAIEILINTYDIEILINDLSSKYVPENMNVKVAARELFIEPEVADADSIDLKWV 299
QY 300 APDEGLISFLVNDGNGFENDRYTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRKET 354
DB 300 EPDEGLVKEFLGDRQFNEVRNKGAKLMSKQAOVQRLDSEFKTLPSTPNATNA-- 357
QY 355 KRKETSDDKTSKAANKKTKAGG 376
DB 358 KKK--AEKAKSANNKAKKATSG 377

RESULT 8

AAW92508 ID AAW92508 standard; Protein; 378 AA.

AC AAW92508;

XX 23-APR-1999 (first entry)

DE Human FEN-1 protein.

XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage.

XX Homo sapiens.
 OS
 XX US874283-A.
 PN
 XX 23-FEB-1999.
 PD
 XX
 XX 30-MAY-1995; 9505-0455968.
 PF
 XX 30-MAY-1995; 9505-0455968.
 PR
 XX (HARR/) HARRINGTON J J.
 PA (HSIEH/) HSIEH C.
 PA (LIEB/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX
 DR WPI; 1999-179985/15.
 DR N-PSDB; AAX02111.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 PS
 XX Disclosure; Fig 5A-B; 58pp; English.
 CC
 CC This sequence represents a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.
 CC
 XX
 XX Sequence 378 AA:

Query Match 51.4%; Score 996; DB 20; Length 378;
 Best Local Similarity 52.4%; Pred. No. 1,6e-83;
 Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;

QY 1 MGKGLTKLLADNAPKAMKEQKFESEYFGKRIADVSMSTIYOFLLIYVGTGNETLTNEAGE 60
 DB 1 MEHGLAKLLADVAPSAIRFENDIKSYFGKRVADSMSTIYOFLLIYVGTGNETLTNEAGE 59
 QY 61 VTSHLQGMFNRTIRLEAGIKPYVYFDGKPPDKKQKQELAKRYSKRDPATKDTLFAVEVGD 120
 DB 60 TTS-LMGVFRYTKR-MENGIKPYVYFDGKPPQKLGKSELAKRSRRRAEAKQLOQAOQAGM 117
 QY 121 KDAIEKLSKRYKVTYRQHNEDCKRLRLKMGVYVVEAPSEAEACALCINDKVPFAVASED 180
 DB 118 EEEVEKTKRYKVTYRQHNEDCKRLRLKMGVYVVEAPSEAEACALCINDKVPFAVASED 177
 QY 181 MDSLTFGAPFLRLHMDPSSKIPVMEFDYAKVLELELTMDQFIDLCITLCCDPCDSIR 240
 DB 178 MDCLTFSPVYRLHMDPSSKIPVMEFDYAKVLELELTMDQFIDLCITLCCDPCDSIR 237
 QY 241 GIGGOTALKLIROHGSIESILENLNDROYIPEDWPYQEARLFEKPNVT-LDIFELKWT 299
 DB 238 GIGGOTALKLIROHGSIESILENLNDROYIPEDWPYQEARLFEKPNVT-LDIFELKWT 297
 QY 300 APDEGLISFLVNDNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATTSAPLRKET 359
 DB 298 EPNEELVKYKFGCEKQFEERIRISGYKRLSKNSOGRLSDFFKPTATTSAPLRKET 356
 QY 360 SDKTSKAAANKKTKAGGKK 379
 DB 357 E--PKGPARKKAKTGAGK 373

RESULT 9
 AAW92506

ID AAW92506 standard. Protein; 382 AA.
 AC AAW92506;
 XX
 XX 23-APR-1999 (first entry)
 DT
 XX
 XX Yeast FEN-1 protein.
 DE
 XX
 XX FEN-1; yeast; flap endonuclease; detection; diagnosis; carcinogen;
 KM neoplasia; antineoplastic agent; cleavage.
 XX
 XX Saccharomyces cerevisiae.
 OS
 XX US874283-A.
 PN
 XX 23-FEB-1999.
 PD
 XX 30-MAY-1995; 9505-0455968.
 PF
 XX 30-MAY-1995; 9505-0455968.
 PR
 XX 30-MAY-1995; 9505-0455968.
 XX
 XX (HARR/) HARRINGTON J J.
 PA (HSIEH/) HSIEH C.
 PA (LIEB/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX
 DR WPI; 1999-179985/15.
 DR N-PSDB; AAX02109.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 PS
 XX Disclosure; Fig 3A; 58pp; English.
 CC
 CC This sequence represents a yeast FEN-1 (flap endonuclease) protein. This
 CC protein is used in a method to isolate novel human FEN-1 proteins for
 CC detecting a pathological condition in a patient, for diagnostic purposes,
 CC for screening for antineoplastic agents and carcinogens, for diagnostic
 CC staging of neoplasia, for producing recombinant flap endonuclease for use
 CC as research or diagnostic reagents, for producing transgenic nonhuman animals
 CC with the novel polypeptides, for producing transgenic nonhuman animals
 CC expressing the novel polypeptides encoded by a transgene. The invention
 CC also provides novel molecular cloning techniques and reagents involving
 CC cleavage of a flap or nick with a flap endonuclease.
 CC
 XX
 XX Sequence 382 AA:

Query Match 48.4%; Score 938; DB 20; Length 382;
 Best Local Similarity 49.5%; Pred. No. 3,8e-78;
 Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGKGLTKLLADNAPKAMKEQKFESEYFGKRIADVSMSTIYOFLLIYVGTGNETLTNEAGE 60
 DB 1 MGKGLTKLLADNAPKAMKEQKFESEYFGKRIADVSMSTIYOFLLIYVGTGNETLTNEAGE 60
 QY 61 VTSHLQGMFNRTIRLEAGIKPYVYFDGKPPDKKQKQELAKRYSKRDPATKDTLFAVEVGD 120
 DB 61 VTSHLQGMFNRTIRLEAGIKPYVYFDGKPPDKKQKQELAKRYSKRDPATKDTLFAVEVGD 120
 QY 121 KDAIEKLSKRYKVTYRQHNEDCKRLRLKMGVYVVEAPSEAEACALCINDKVPFAVASED 180
 DB 121 K--MQOERLVYVSKHEHEAQKLLGTMGIPYIAPTEAEVQCAELAKKGVAAASED 177
 QY 181 MDSLTFGAPFLRLHMDPSSKIPVMEFDYAKVLELELTMDQFIDLCITLCCDPCDSIR 240
 DB 178 MDCLTFSPVYRLHMDPSSKIPVMEFDYAKVLELELTMDQFIDLCITLCCDPCDSIR 237
 QY 241 GIGGOTALKLIROHGSIESILENLNDROYIPEDWPYQEARLFEKPNVTLDIPE 295
 DB 238 GIGGOTALKLIROHGSIESILENLNDROYIPEDWPYQEARLFEKPNVTLDIPE 295
 QY 296 --LKWAPDEGLISFLVNDNGFNEDRYTKAIEKISAKNKSOGRLSEFFKPTATT--- 350

DB 297 INKMSPPKREKLEIYLCDDKFKSEERYKSGISRLKRLKSGIGRLDGFQVVPKTRKQ 356
QY 351 -SAPLKRKETSDDTKSKAANKTKR 373
DB 357 LAAMAKRAOENKRLNK-NKNKYTK 379

RESULT 10
ABG19545
ID ABG19545 standard; Protein: 373 AA.
AC ABG19545:
XX
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE Novel human diagnostic protein #19536.
XX
XX
KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX
PA (HYSE-) HYSEQ INC.
PI Dmanac RT, Liu C, Tang YT;
DR WPI: 2001-639362/73.
DR N-PSDB: AAS83732.
XX
XX
PT New isolated polynucleotide and encoded polypeptides; useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX
PS Claim 20; SEQ ID No 49904; 103pp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX
SQ Sequence 373 AA:
Query Match 36.1%; Score 699.5; DB 22; Length 373;
Best Local Similarity 41.1%; Pred. No. 4.1e-56;
Matches 155; Conservative 68; Mismatches 105; Indels 49; Gaps 7;

QY 1 MGKGTFLTLADNMPKAMEKQKFESEYGRKTAVDASMSIYQFLIVGRGMETLTNAGE 60
DB 36 MGIGLAKLADVAPSAIRENDIKSYFGRKVALDASMSITHFLLAV-RGGDVLQNEEG 94
QY 61 VTSHLQGFNRTIRLENGIKPVYVFDGKPPDMKQELAKRYSKRDATDGLTEAVEGD 120
DB 95 TTSHLMGFEYRTIRKENGIKPIY-----KSGELAKSSELRAFAKQLOQAAGA 145
QY 121 KDAIEKSKRTYKPYVRQNEDECKRLRLMLGVPVVEASEAEACALCIDKXFAVASE 180
DB 146 EGVVEKFTKWLKVKYKQNDCKHLSLMDGIPYDAPLEASCAALYKAGKXYPVATED 205
QY 181 MDLTFGAPRFLRLMDSSKRIIPVMEFDVAKVLELELTMDQFIDLCILCGCDYDSIK 240
DB 206 MDCFTFGSPVLMQHLTASS-----DYCKSIW 231
QY 241 GIGGOTALKLIRHSGSISILENKKDROYTEPDMPYQEARLFEKPNVT-LDPELXMT 299
DB 232 STGPRAYDLQKHRSIEIYQRLDPNKYPLPENRLHKEAYQLFLEPEVLGPESVELKMS 291
QY 300 ABDEGLISFLVKDNGFNEDRVYKAIEKIKSNKSSQGRLESFFKPTATTSAPLKRKE- 358
DB 292 ENEBELVKFVCGEKQFSEERHSGVKRLRTEKG-GGEPRGPDWTP--GRASPTKRPDH 348
QY 359 TSDKTSKAANKTKKAG 375
DB 349 QEDQHTSRKSSSESHXG 365

RESULT 11
AAM24216
ID AAM24216 standard; Protein: 340 AA.
XX
AC AAM24216;
XX
DT 14-APR-1998 (first entry)
XX
DE Pyrococcus furiosus FEN-1 endonuclease.
XX
XX
KM Nucleic acid cleavage: 5' nuclease; DNA cleavage; RNA cleavage;
KM invader directed cleavage; FEN-1; endonuclease.
XX
XX
OS Pyrococcus furiosus.
XX
XX
PN WO9727214-A1.
XX
PD 31-JUL-1997.
XX
PF 22-JAN-1997; 97WO-US01072.
XX
PR 02-DEC-1996; 96US-0759038.
PR 24-JAN-1996; 96US-0599491.
PR 12-JUL-1996; 96US-0682853.
PR 29-NOV-1996; 96US-0756386.
PR 02-DEC-1996; 96US-0758314.
XX
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichiev VI;
PI Olive DM, Prudent JR;
XX
XX
DR WPI: 1997-393613/36.
DR N-PSDB: AAT76685.
XX
XX
PT Thermostable structure-specific nuclease(s) - used for detection and
PT characterisation of nucleic acid sequences and variations in nucleic
PT acid sequences
XX
XX
XX Example 28; Page 283-285; 457pp; English.
XX
CC This sequence comprises Pyrococcus furiosus (Pfu) FEN-1
CC endonuclease. Large-scale production of the enzyme was performed

using *E. coli* host cells transformed with a vector carrying the *Pfu* FEN-1 coding sequence (see AAT7668). *Pfu* FEN-1 is a thermostable CC enzyme. It can be used in novel methods for the detection and CC characterisation of nucleic acid sequences and variations in CC nucleic acid sequences.

XX Sequence 340 AA:

Query Match 33.4%; Score 648.5; DB 18; Length 340;
Best Local Similarity 41.3%; Pred. No. 1.8e-51;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGKRTAVDASMSIYQFLIVGRGMEITLNAGEVTSHLOGMNRITRLLEA 78
D 12 KEITLENLYGKKTAIDALNAIYQFLSTIRKQDTPLMDSKGRITSHLSGLFYRTINLMEA 71
QY 79 GIKRVYVFDKPPDMKQELAKRYRKDDATKDLTEAVGVGKDAIEKLSKRTVAVTRQH 138
D 72 GIKRVYVFDKPPDMKQELAKRYRKDDATKDLTEAVGVGKDAIEKLSKRTVAVTRQH 131
QY 139 NEDCKRLRLMGYPVVEAPSEAEACALCINDKVFVASEDMDSITFGAPRFLHMDP 198
D 132 IEDAKKLTLELMGIPYQAPSEGEQAAYMAKGSYASASQDYDSLFGAPRLVNLITTT 191
QY 199 SSKRIP-----VMEFDVAKVLELELTMDOFDICICGGDY-CDSTKIGIG 244
D 192 GKRLPGKNYVVEIKPELITLLE-----EVLEKLTIREKLEIALIVGTDYNGGKIGL 247
QY 245 QTALKLIRQHGSIESTILENLNKD--RYQIPEDWPYQEARLFKEPNVTLDIPELKWAP 301
D 248 KKALEIYRHH-----SKDPLAKFKQSDVDLYAIKEFLNPVT-DNYNLWMDP 295
QY 302 DEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 345
D 296 DEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 339

RESULT 12
AAM79970
ID AAM79970 standard; Protein; 340 AA.

AC AAM79970;

DT 02-FEB-1999 (first entry)

XX Pyrococcus furiosus FEN-1 endonuclease.

XX Nucleic acid detection; multiple sequential invasive cleavage;

KW FEN-1; endonuclease; nuclease.

XX Pyrococcus furiosus.

PN W09842873-A1.

PD 01-OCT-1998.

PE 24-MAR-1998; 98WO-US05809.

PR 24-MAR-1997; 97US-0823516.

PA (THIR-) THIRD WAVE TECHNOLOGIES INC.

PI Brow MAD, Hall JG, Kwiatkowski RM, Lyamlichev VI;

DR Mast AL, Vavra SH;

DR WPI; 1998-557036/47.

DR N-PSDB; AAV65840.

PT Detecting target nucleic acid by sequence-specific cleavage of

PT complex with two specific oligonucleotides - used to detect

XX cytomagalovirus DNA

PS Example 28b; Page 316-317; 524pp; English.

XX This is the amino acid sequence of FEN-1 endonuclease of CC *Pyrococcus furiosus* (Pfu). FEN-1 DNA (see AAV65840) has been CC ligated into vector for expression in *E. coli* cells. The invention CC relates to means for the detection and characterisation of nucleic CC acid sequences, and variations in nucleic acid sequences. It also CC relates to methods for forming a nucleic acid cleavage structure on CC a target sequence and cleaving this structure in a site-specific CC manner, preferably using a thermostable structure-specific nuclease CC such as FEN-1. cleavage of the cleavage structure by the CC nuclease indicates the presence of specific nucleic acid sequences CC or specific variants. The invention further relates to methods for CC the separation of nucleic acid molecules based on charge, methods CC for the detection of non-target cleavage products via the formation CC of a complete and activated protein binding region, and methods for CC the detection of nucleic acid from various viruses (e.g. human CC cytomagalovirus) in a sample. The method amplifies the detection CC molecule rather than the target itself, is less subject to CC contamination than exponential amplification processes, and allows CC many targets to be analysed in a single reaction.

XX Sequence 340 AA:

Query Match 33.4%; Score 648.5; DB 19; Length 340;
Best Local Similarity 41.3%; Pred. No. 1.8e-51;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGKRTAVDASMSIYQFLIVGRGMEITLNAGEVTSHLOGMNRITRLLEA 78
D 12 KEITLENLYGKKTAIDALNAIYQFLSTIRKQDTPLMDSKGRITSHLSGLFYRTINLMEA 71
QY 79 GIKRVYVFDKPPDMKQELAKRYRKDDATKDLTEAVGVGKDAIEKLSKRTVAVTRQH 138
D 72 GIKRVYVFDKPPDMKQELAKRYRKDDATKDLTEAVGVGKDAIEKLSKRTVAVTRQH 131
QY 139 NEDCKRLRLMGYPVVEAPSEAEACALCINDKVFVASEDMDSITFGAPRFLHMDP 198
D 132 IEDAKKLTLELMGIPYQAPSEGEQAAYMAKGSYASASQDYDSLFGAPRLVNLITTT 191
QY 199 SSKRIP-----VMEFDVAKVLELELTMDOFDICICGGDY-CDSTKIGIG 244
D 192 GKRLPGKNYVVEIKPELITLLE-----EVLEKLTIREKLEIALIVGTDYNGGKIGL 247
QY 245 QTALKLIRQHGSIESTILENLNKD--RYQIPEDWPYQEARLFKEPNVTLDIPELKWAP 301
D 248 KKALEIYRHH-----SKDPLAKFKQSDVDLYAIKEFLNPVT-DNYNLWMDP 295
QY 302 DEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 345
D 296 DEGLISFLVKGNGFNEDRYTKAIEKISAKNKSOGRLSEFPK 339

RESULT 13
AAM59940
ID AAM59940 standard; Protein; 340 AA.

AC AAM59940;

DT 21-DEC-1998 (first entry)

XX Amino acid sequence of the *Pfu* FEN-1 endonuclease.

KW *Mja* FEN-1 endonuclease; Tag gene; structure-specific nuclease;

KW mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;

KW hepatitis C virus; HCV; thermostable.

XX Pyrococcus furiosus.

PN W09823774-A1.

PD 04-JUN-1998.

PE 26-NOV-1997; 97WO-US21783.

```

XX 02-DEC-1996: 96US-0758314.
PR 29-NOV-1996: 96US-0757653.
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
PA Kaiser MW, Lyamichev VI, Lyamicheva N;
PI WPI: 1998-322748/28.
DR N-PSDB: AAV53951.
XX Thermostable structure-specific nuclease(s) derived from mutant DNA
PT polymerase(s) - useful for detecting mutant allele(s) or strains of
PT microorganisms
XX Example 29: Pages 280-281; 472pp; English.
PS This is the amino acid sequence encoding the pfu FEN-1 endonuclease,
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).
XX Sequence 340 AA:
SQ
Query Match 33.4%; Score 648.5; DB 19; Length 340;
Best Local Similarity 41.3%; Pred. No. 1.8e-51;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;
QY 19 KEQKESYFGKRIADVADSMSTYOFILVGRGMEITLTNEAGEVTSHQGMFNRTIRLEA 78
DB 12 KEIENLNYGKRIADLALNAYOFLSTIRQDGTPLMDSKGRITSHSGLFYRTINLMEA 71
QY 79 GIKPYVFDGKPPDMKKQOELAKRYSKRDATKDLTEAVEVGDKAIEKLSKRTVYTRQH 138
DB 72 GIKPYVFDGPPPEPKKELEKREAREEAEKREAEKLEIEBKRYAKQATRYNEML 131
QY 139 NEDCKRLRLMGVAVVPAPEAECAALCINDKVFVAVASEDMDSLTFGAPFLRLMDP 198
DB 132 IEDAKKLELMGIPYVQAPSEGAQAYMAKSYASASQDYDSLTCGAPRLVNLTT 191
QY 199 SSKKIP-----VMEFVAKVLELELTMDQFIDLCICGCDY-CDSTKGTGG 244
DB 192 GRRKLPGKNVYEIKPELITL---EVLEKELTRKELIETALIVGTDYNGGKIGL 247
QY 245 OTALKLIRHGSIESILENLNKD---RYQIPEDMPYQEARLRFKEPNVTLDIPELKTAP 301
DB 248 KKALEIVRH-----SKDPLAKFQKQSDVDLYAIKEFLPNPVT-DNYNLVWRDP 295
QY 302 DEEGILSPFYKDNCFNEDRYTAKETIKSAKNNKSOGRLSEFPK 345
DB 296 DEEGILKFLCDHDSERKANKGLERLKAIKSGKOSTLESMPK 339
RESULT 14
AAW59953
ID AAW59953 standard; Protein: 326 AA.
XX
AC AAW59953:
XX
DT 22-DEC-1998 (first entry)
XX
DE Amino acid sequence of the structure specific nuclease 5.
XX
KW Tag mutant gene; thermostable; nuclease; mutant; DNA polymerase;
XX bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
XX
OS Synthetic.

```

```

XX W09823774-A1.
XX
XX 04-JUN-1998.
XX
XX 26-NOV-1997: 97WO-US21783.
XX
XX 02-DEC-1996: 96US-0758314.
PR 29-NOV-1996: 96US-0757653.
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
PA Kaiser MW, Lyamichev VI, Lyamicheva N;
PI WPI: 1998-322748/28.
DR Thermostable structure-specific nuclease(s) derived from mutant DNA
PT polymerase(s) - useful for detecting mutant allele(s) or strains of
PT microorganisms
XX Claim 1: Pages 331-332; 472pp; English.
PS This is the amino acid sequence of a structure specific nuclease
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).
XX Sequence 326 AA:
SQ
Query Match 33.4%; Score 647; DB 19; Length 326;
Best Local Similarity 42.6%; Pred. No. 2.4e-51;
Matches 140; Conservative 70; Mismatches 103; Indels 16; Gaps 7;
QY 19 KEQKESYFGKRIADVADSMSTYOFILVGRGMEITLTNEAGEVTSHQGMFNRTIRLEA 78
DB 12 KEIENLNYGKRIADLALNAYOFLSTIRQDGTPLMDSKGRITSHSGLFYRTINLMEA 71
QY 79 GIKPYVFDGKPPDMKKQOELAKRYSKRDATKDLTEAVEVGDKAIEKLSKRTVYTRQH 138
DB 72 GIKPYVFDGPPPEPKKELEKREAREEAEKREAEKLEIEBKRYAKQATRYNEML 131
QY 139 NEDCKRLRLMGVAVVPAPEAECAALCINDKVFVAVASEDMDSLTFGAPFLRLMDP 198
DB 132 IEDAKKLELMGIPYVQAPSEGAQAYMAKSYASASQDYDALYGAAPRVVRNL-T 189
QY 199 SSKKIPVMEFVAKVLELELTMDQFIDLCICGCDY-CDSTKGTGGOTLKLIRHGS 257
DB 190 TTKEMPEL-TEINLEVEDLRLISDLDLITAFMGTDYNGGKIGRRAYELVRS-GVA 247
QY 258 ESILENLKDRYQIPEDMPYQEARLRFKEPNVTLDIPELKTAPDEGLTISFLKONGEN 317
DB 248 KDVLKK-----EVEYYDEIKRIKFEKPYT-DNYSLSLKLDPKGGIITFLVDENDN 297
QY 318 EDRYTKAIEKIKS-AKNNKSOGRLSEFPK 345
DB 298 YDRVKRHVDKLYNLIAWKTKQTLDAMFK 326
RESULT 15
AAW59951
ID AAW59951 standard; Protein: 332 AA.
XX
AC AAW59951:
XX
DT 22-DEC-1998 (first entry)
XX
DE Amino acid sequence of the structure specific nuclease 3.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:27:44 ; Search time 16 Seconds
(without alignments)
578.581 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939

Sequence: 1 MGKGLTKRLADNAPKAKE.....SDKTSKAAANKTKAGGKKK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA:*

1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	379	4	US-09-426-557-2
2	1939	100.0	379	4	US-09-426-557-6
3	1933	99.7	379	4	US-09-426-557-4
4	1933	99.7	379	4	US-09-426-557-8
5	1032	53.2	380	2	US-08-823-516-137
6	1032	53.2	380	4	US-09-426-557-10
7	1031	53.2	380	2	US-08-435-968E-1
8	1014	52.3	378	2	US-08-823-516-138
9	1007.5	52.0	378	2	US-08-455-968E-3
10	996	51.4	382	2	US-08-455-968E-5
11	938	48.4	382	2	US-08-823-516-119
12	938	48.4	382	2	US-08-823-516-119
13	648.5	33.4	340	2	US-08-757-653-176
14	648.5	33.4	340	2	US-08-823-516-79
15	648.5	33.4	340	2	US-08-823-516-116
16	648.5	33.4	340	2	US-08-759-038-115
17	648.5	33.4	340	3	US-08-758-314-115
18	632.5	32.6	343	4	US-09-146-319-2
19	632.5	32.6	343	4	US-09-175-973-2
20	560.5	28.9	326	2	US-08-757-653-112
21	560.5	28.9	326	2	US-08-823-516-75
22	560.5	28.9	326	2	US-08-823-516-135
23	560.5	28.9	326	3	US-08-759-038-111
24	560.5	28.9	326	3	US-08-758-314-111
25	364	18.8	387	2	US-08-823-516-140
26	352.5	18.2	488	2	US-08-823-516-141
27	340.5	17.6	386	2	US-08-455-968E-7

28	340	17.5	543	2	US-08-823-516-143	Sequence 143, App
29	334.5	17.3	530	2	US-08-823-516-142	Sequence 142, App
30	313.5	16.2	527	2	US-08-823-516-144	Sequence 144, App
31	205	10.6	872	1	US-08-766-014-2	Sequence 2, Appl1
32	184.5	9.5	291	4	US-09-105-697-6	Sequence 6, Appl1
33	173.5	8.9	834	5	PCT-US95-14418-2	Sequence 2, Appl1
34	173.5	8.9	834	5	PCT-US95-15327-2	Sequence 2, Appl1
35	173	8.9	548	2	US-08-484-956-86	Sequence 86, Appl
36	173	8.9	548	2	US-08-757-653-86	Sequence 86, Appl
37	173	8.9	695	2	US-08-484-956-87	Sequence 87, Appl
38	173	8.9	695	2	US-08-757-653-87	Sequence 87, Appl
39	173	8.9	810	4	US-09-587-856-2	Sequence 2, Appl1
40	173	8.9	810	4	US-09-777-537-2	Sequence 2, Appl1
41	173	8.9	810	4	US-09-777-538-2	Sequence 2, Appl1
42	173	8.9	832	1	US-07-977-434-2	Sequence 2, Appl1
43	173	8.9	832	1	US-08-156-020-2	Sequence 2, Appl1
44	173	8.9	832	1	US-08-073-384C-4	Sequence 4, Appl1
45	173	8.9	832	1	US-08-254-359A-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-426-557-2
; Sequence 2, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE OF INVENTION: Theroof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-2

Query Match	Score	1939;	DB 4;	Length	379;
Best Local Similarity	100.0%;	Pred. No. 1e-183;			
Matches 379;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY 1	MGKGLTKRLADNAPKAKEKFESEYFGKRIAYDASMSIYQFLIVGRTGMETLTNAGE	60			
DB 1	MGKGLTKRLADNAPKAKEKFESEYFGKRIAYDASMSIYQFLIVGRTGMETLTNAGE	60			
QY 61	VNSHLOGMNRITRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDDATYDLTEAVEVGD	120			
DB 61	VNSHLOGMNRITRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDDATYDLTEAVEVGD	120			
QY 121	KDAIEKLSKRYKVTROHNECDKRLRLMGVYVVEAPSEAECAALCINDKFAVASED	180			
DB 121	KDAIEKLSKRYKVTROHNECDKRLRLMGVYVVEAPSEAECAALCINDKFAVASED	180			
QY 181	MSDLTGARFLRLHMDPSKKIPYMEFDYAKVLELELTMDQITDLCILCGDYCDSIK	240			
DB 181	MSDLTGARFLRLHMDPSKKIPYMEFDYAKVLELELTMDQITDLCILCGDYCDSIK	240			
QY 241	GIGGOTATLRIQHGSIIESIENLNKRDYQIPEDMPYQEARLRFKEPNVTLDIPELKWTA	300			
DB 241	GIGGOTATLRIQHGSIIESIENLNKRDYQIPEDMPYQEARLRFKEPNVTLDIPELKWTA	300			
QY 301	PDEEGILSLVYNDGNFNEEDRYTKALEIKSAKNSQGRLEFFKPPATYSAPLKRKETS	360			
DB 301	PDEEGILSLVYNDGNFNEEDRYTKALEIKSAKNSQGRLEFFKPPATYSAPLKRKETS	360			
QY 361	DKTSKAAANKTKAGGKKK 379				

Db 361 DKTSAANKKTAGGKK 379

RESULT 2

```
US-09-426-557-6
; Sequence 6, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; EARLIER FILING DATE: 1999-10-22
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-6
```

Query Match

```
Best Local Similarity 100.0%; Score 1939; DB 4; Length 379;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVGRGTMETLTNEAGE 60
Db 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVGRGTMETLTNEAGE 60
Qy 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
Db 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
Qy 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180
Db 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180
Qy 181 MSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGGCDYCSIK 240
Db 181 MSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGGCDYCSIK 240
Qy 241 GIGGQTAALKIRHOGSIESILENLKDRQIPEDPWYQEARRLKPEPNVTLDIPELKWTA 300
Db 241 GIGGQTAALKIRHOGSIESILENLKDRQIPEDPWYQEARRLKPEPNVTLDIPELKWTA 300
Qy 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
Db 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
Qy 361 DKTSAANKKTAGGKK 379
Db 361 DKTSAANKKTAGGKK 379
```

RESULT 3

```
US-09-426-557-4
; Sequence 4, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; EARLIER FILING DATE: 1999-10-22
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-4
```

Query Match

```
Best Local Similarity 99.7%; Score 1933; DB 4; Length 379;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVGRGTMETLTNEAGE 60
Db 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVGRGTMETLTNEAGE 60
Qy 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
Db 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
Qy 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180
Db 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180
Qy 181 MSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGGCDYCSIK 240
Db 181 MSLTFGAPREFLRHLMDDPSKKIPVMEFDVAKVLELELTMDQFLDLCILGGCDYCSIK 240
Qy 241 GIGGQTAALKIRHOGSIESILENLKDRQIPEDPWYQEARRLKPEPNVTLDIPELKWTA 300
Db 241 GIGGQTAALKIRHOGSIESILENLKDRQIPEDPWYQEARRLKPEPNVTLDIPELKWTA 300
Qy 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
Db 301 PDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRKETS 360
Qy 361 DKTSAANKKTAGGKK 379
Db 361 DKTSAANKKTAGGKK 379
```

RESULT 4

```
US-09-426-557-8
; Sequence 8, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; EARLIER FILING DATE: 1999-10-22
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-8
```

Query Match

```
Best Local Similarity 99.7%; Score 1933; DB 4; Length 379;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVGRGTMETLTNEAGE 60
Db 1 MGKIGLTKLLADNAPKAMEQKFESEYFGKRIAVDASMSIYOFLLVGRGTMETLTNEAGE 60
Qy 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
Db 61 VTSHLOGMFNRTIRLLEAGIKPVYFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
Qy 121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVYVVEAPSEAECAALCINDKVFVASED 180
```

```

;=====
Db 121 KDAIEKLSKRTVYKQHNEDCKRLRLMGVYVPEAPSEAEBCALCINDKVFAYASED 180
QY 181 MSLTFEGAPRFLRLHMDSSKKIPMEFDVAKYLELELIMDOFIDLCICGDCYCSIK 240
Db 181 KOSLTFEGAPRFLRLHMDSSKKIPMEFDVAKYLELELIMDOFIDLCICGDCYCSIK 240
QY 241 GIGGOTALKLIRHGSISLLENLNKDRYQIPEDMPYQEARRLKEPNVTLDIPELKWTA 300
Db 241 GIGGOTALKLIRHGSISLLENLNKDRYQIPEDMPYQEARRLKEPNVTLDIPELKWTA 300
QY 301 PDEELISFLVNDGNFNDRTYKALEIKSKAKNSGGRLESPFKPATTSAPLKRKETS 360
Db 301 PDEELISFLVNDGNFNDRTYKALEIKSKAKNSGGRLESPFKPATTSAPLKRKETS 360
QY 361 DKTSAANKKTKAGGKK 379
Db 361 DKTSAANKKTKAGGKK 379

RESULT 5
US-08-823-516-137
; Sequence 137, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FGRS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338

```

```

; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-823-516-137

Query Match 53.2%; Score 1032; DB 2; Length 380;
Best Local Similarity 53.8%; Pred. No. 6,1e-94;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGKGLTKLADNAPKAMKEQKFESEYGRKIAVDASMSIYQFLIVYGRGMEITLNEAGE 60
Db 1 MGIGGLTKLADVAPSAIRENDIKSYGRKVAIDASMSIYQFLIAV-ROGDVILQNEE 59
QY 61 VTSHOGMFNTIRLLEAGIKPVYFPDGKPPDKKQFLARYSKRDDATKDLTEAVEVD 120
Db 60 TTSILMGHFTIRKMEGILPVYFDGKPPDLKSGELARSRERAEKQLOQAAGA 119
QY 121 KDAIEKLSKRTVYKQHNEDCKRLRLMGVYVPEAPSEAEBCALCINDKVFAYASED 180
Db 120 EOEVEKFTKRLVYKQHNEDCKRLSLMGIPYLDAPSEAEBCALVYKGYAATED 179
QY 181 MSLTFEGAPRFLRLHMDSSKKIPMEFDVAKYLELELIMDOFIDLCICGDCYCSIK 240
Db 180 MCLTFEGAPRFLRLHMDSSKKIPMEFDVAKYLELELIMDOFIDLCIGSDYCSIR 239
QY 241 GIGGOTALKLIRHGSISLLENLNKDRYQIPEDMPYQEARRLKEPNVTLDIPELKW 298
Db 240 GIGPRAVDLIQKHSEIYRILDPKPKYPPEWMLRKEAQLFLEPEV-LDESEVELKW 298
QY 299 TAPDEGLISFLVNDGNFNDRTYKALEIKSKAKNSGGRLESPFKPATTSAPLKRKE 358
Db 299 SEPNEELIKFWGEGKQPSSEIRNSGVKRLSKSRQSGRLDDFRVYGLSS-AKKE 357
QY 359 TSDK--TSAANKTKYAGGKK 379
Db 358 PEPKSTKKKA---KTGAAGKK 377

RESULT 6
US-09-426-557-10
; Sequence 10, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-426-557-10

Query Match 53.2%; Score 1032; DB 4; Length 380;
Best Local Similarity 53.8%; Pred. No. 6,1e-94;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY 1 MGKGLTKLADNAPKAMKEQKFESEYGRKIAVDASMSIYQFLIVYGRGMEITLNEAGE 60
Db 1 MGIGGLTKLADVAPSAIRENDIKSYGRKVAIDASMSIYQFLIAV-ROGDVILQNEE 59
QY 61 VTSHOGMFNTIRLLEAGIKPVYFPDGKPPDKKQFLARYSKRDDATKDLTEAVEVD 120

```

Dh 60 TTSHLGMFRTIRRMENGIRKPYVFDGKRPOLKSGELAKRSERAEAKOLOQOAGAGA 119
QY 121 KDAIEKLSKRTVAVTROHNEDECKRLRLMGVAVVPAEPAEBCALCINDKRVAVASED 180
Db 120 EQEVEKFTKLAVVTKOHNEDECKHLSLGMGIPYLDAPSEAEASCAALVKGKVVAAATED 179
QY 181 MDSLTFGAPRFLRHLMDPSSKRIPVMEFDVAVLELELTMDQFDLCLTLCGCGYCSIRK 240
Db 180 MDCLTFGSPVLMRHLTASAEAKKLPIQEFHLSRLLOELGLNGQFVDFCLTLGSDYCESIR 239
QY 241 GIGGOTALKLIRHOGSIESTELENKKDRYQIPEDMVPOEARLFEKPEVYTD--IPELKW 298
Db 240 GIGPKRVDLIQKHKSIEEIVRLDPNKYPVENMLKHAHOLELEPEV-LDPESVELKW 298
QY 299 TAPDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRE 358
Db 299 SEPNEELIKFMGEGKQFSEERIRSGVKRLSKROGSTQGRLDPFKVTGSLS--AKRKE 357
QY 359 TSDK--TSKAANKTKRAGKKK 379
Db 358 PEPKSTKKKA--KTGAAGKFK 377

RESULT 7
US-08-455-968E-1
; Sequence 1, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-455-968E-1

Query Match 53.2%; Score 1031; DB 2; Length 380;
Best Local Similarity 53.8%; Pred. No. 7.7e-94;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;
QY 1 MGIGLTKLADNAPKAMEKQEFSEYFGKRIKIVDASMSIYQELIYVAGTETLTNAGE 60
Db 1 MGIGLTKLADNAPKAMEKQEFSEYFGKRIKIVDASMSIYQELIYVAGTETLTNAGE 59

QY 61 VTSHLOGMNFRTIRLLEAGIKPYVYFDGKPPDMKKOELAKRSRDKDXTKDLTEAVEGD 120
Db 60 TTSHLGMFRTIRRMENGIRKPYVFDGKRPOLKSGELAKRSERAEAKOLOQOAGAGA 119
QY 121 KDAIEKLSKRTVAVTROHNEDECKRLRLMGVAVVPAEPAEBCALCINDKRVAVASED 180
Db 120 EQEVEKFTKLAVVTKOHNEDECKHLSLGMGIPYLDAPSEAEASCAALVKGKVVAAATED 179
QY 181 MDSLTFGAPRFLRHLMDPSSKRIPVMEFDVAVLELELTMDQFDLCLTLCGCGYCSIRK 240
Db 180 MDCLTFGSPVLMRHLTASAEAKKLPIQEFHLSRLLOELGLNGQFVDFCLTLGSDYCESIR 239
QY 241 GIGGOTALKLIRHOGSIESTELENKKDRYQIPEDMVPOEARLFEKPEVYTD--IPELKW 298
Db 240 GIGPKRVDLIQKHKSIEEIVRLDPNKYPVENMLKHAHOLELEPEV-LDPESVELKW 298
QY 299 TAPDEGLISFLVKNDFNEDRVTKAIEKISAKNKSOGRLSEFFKPTATTSAPLKRE 358
Db 299 SEPNEELIKFMGEGKQFSEERIRSGVKRLSKROGSTQGRLDPFKVTGSLS--AKRKE 357
QY 359 TSDK--TSKAANKTKRAGKKK 379
Db 358 PEPKSTKKKA--KTGAAGKFK 377

RESULT 8
US-08-823-516-138
; Sequence 138, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichay, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brox, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PCT/US97/01072
; APPLICATION NUMBER: 435
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 138:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-823-516-138

Query Match 52.3%; Score 1014; DB 2; Length 378;
Best Local Similarity 53.2%; Pred. No. 3,7e-92;
Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGKGLTKLADNAPKAKKEQKESYFGKRTIADVDSMSIYQFLIVGRTGMEITLNEAGE 60
DB 1 MGIHGLAKLIADVAPSAIRENDIKSYFGKVAIDASMSIYQFLIAY-RQGSDVLDONEEGE 59
QY 61 VSHLOGMNFRTIRLEAGIKPVYVFDGKPPDMKROELAKRYSKRDATKDLTEAVEVGD 120
DB 60 TTS-LMGMEYRTIR-MENGIRKPVYVFDGKPPOLKSGELAKRSERAEKQLOQAQAGM 117
QY 121 KDAIEKLSKRYKYTRQNEDECKRLRLMGVYVVEAESEAEACALCINDKYFAVASED 180
DB 118 EEEVEFTKRLVKTQKQNDCKHLISLMGIPYLDAPSEAEASCAALAKAGKYAAATED 177
QY 181 MSLTFGARFLRLHMDSSKKIPMEFDVAKVLEELTMDQFIDLCILGDCYDCE 240
DB 178 MCLTFGSPVLMRHLTASEAKKLPIQEPHLSRVLDQELGNDQVDCILGSDYCESIR 237
QY 241 GIGGQATALKLRQHSISILENKNDRYQIPEDMPYQEARLKEPNVT-LDIPELKWT 299
DB 238 GIGARAVDLQKHRSIEIVRLDPSKYVPVENMLHKEAQLLEPEVVDPESEVELKWS 297
QY 300 APDEGLISFLVKDGFNEDRYTAKIEKISAKNKSQGRLESEFKRTATTSAPLKRKT 359
DB 298 EPNEELVAFMGCEKQFSEIRSGVKRLSKRSGSTQGRLDQFFKYTGLSS-ARKKEP 356
QY 360 SDKTSKAANKTKAGGKK 379
DB 357 E--PRGPAKKRKTGAGK 373

RESULT 9
US-08-455-968E-3
Sequence 3, Application US/08455968E

PATENT No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0300
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-455-968E-3

Query Match 52.0%; Score 1007.5; DB 2; Length 377;
Best Local Similarity 53.4%; Pred. No. 1,6e-91;
Matches 203; Conservative 68; Mismatches 100; Indels 9; Gaps 7;

QY 1 MGKGLTKLADNAPKAKKEQKESYFGKRTIADVDSMSIYQFLIVGRTGMEITLNEAGE 60
DB 1 MGIHGLAKLIADVAPSAIRENDIKSYFGKVAIDASMSIYQFLIAY-RQGSDVLDONEEGE 59
QY 61 VSHLOGMNFRTIRLEAGIKPVYVFDGKPPDMKROELAKRYSKRDATKDLTEAVEVGD 120
DB 60 TTS-LMGMEYRTIR-MENGIRKPVYVFDGKPPOLKSGELAKRSERAEKQLOQAQAGM 116
QY 121 KDAIEKLSKRYKYTRQNEDECKRLRLMGVYVVEAESEAEACALCINDKYFAVASED 180
DB 117 MEVEEFKRLVKTQKQNDCKHLISLMGIPYLDAPSEAEASCAALAKAGKYAAATED 176
QY 181 MSLTFGARFLRLHMDSSKKIPMEFDVAKVLEELTMDQFIDLCILGDCYDCE 240
DB 177 MCLTFGSPVLMRHLTASEAKKLPIQEPHLSRVLDQELGNDQVDCILGSDYCESIR 236
QY 241 GIGGQATALKLRQHSISILENKNDRYQIPEDMPYQEARLKEPNVT-LDIPELKWT 299
DB 237 GIGARAVDLQKHRSIEIVRLDPSKYVPVENMLHKEAQLLEPEVVDPESEVELKWS 296
QY 300 APDEGLISFLVKDGFNEDRYTAKIEKISAKNKSQGRLESEFKRTATTSAPLKRKT 359
DB 297 EPNEELVAFMGCEKQFSEIRSGVKRLSKRSGSTQGRLDQFFKYTGLSS-ARKKEP 355
QY 360 SDKTSKAANKTKAGGKK 379
DB 356 EPKGS--AKKRAKTGAGK 372

RESULT 10
US-08-455-968E-10
Sequence 10, Application US/08455968E

PATENT No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E

```

; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-455-968E-10

Query Match      51.4%; Score 996; DB 2; Length 378;
Best Local Similarity 52.4%; Pred. No. 2.2e-90;
Matches 199; Conservative 70; Mismatches 103; Indels 8; Gaps 6;

QY 1 MGIGLTKLADNAPKAMKEQKFESEYFGRIAVDASMSIYQFLIYVGRGMEITLNEAGE 60
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MEHGLAKLIADYAPSAIRENDIKSFGRKVAIDASMSIYQFLIYV-RGGGDVLQNEGE 59
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 VTSHLQGMENRTILLEGIRPVYFDGKPPDMKQOELAKRYSKRDATKDLTEAVEVGD 120
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 TTS-LMGMEYRFRW-MENGIKPVYVFDGKPPDLKSGELAKRSERAEKOLOAOQAGM 117
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KDALEKLSRTKYVTRQHNEDCKRLRLMGVNVPEAPSEAEACALCINDVFAVASED 180
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 118 EEEVEFTKRLVAVYQHNDECKHLSDMGIPYLDAPSAEASCAALAKAGVYAATED 177
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 MDSLTFGAPRFLRLHMDPSSKRIPVMEFDYAKVLELELMDQIDLCILGCDYCSIR 240
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 MDCLTFGSPVLMRLHASEAKKLPIDFHLISRYLOELGINOQFVLDICILGSDYCESIR 237
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 GIGGOTAKLIRQHSISTELENKDRQIPEDMPYQEARLFEKPNVNT-LDIELELWT 299
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 GIGAKRRAVDLQHKSTIEELVRLDSKTFVPMHLKKEQOLFLEPYVDESEVELKMS 297
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 APDEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGRLSEFFKPTATTSAPLRKET 359
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 EPNEELVKNMGCEKQFEERIRSGVKRLSKSGOSTGRIDDFEYVTSLSL-AKRKEP 356
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 SDKTSRAANKTKAGGKK 379
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 E---PKGPARKKAKTGGAK 373
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-08-455-968E-5
; Sequence 5, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-455-968E-5

Query Match      48.4%; Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 1.2e-84;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGIGLTKLADNAPKAMKEQKFESEYFGRIAVDASMSIYQFLIYVGRGMEITLNEAGE 60
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MGIGLNAIIEHVPASAIRKSDIKSFGRKVAIDASMSIYQFLIYVAVROQDGGQITNEAGE 60
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 VTSHLQGMENRTILLEGIRPVYFDGKPPDMKQOELAKRYSKRDATKDLTEAVEVGD 120
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 TTSHLQGMENRTILLEGIRPVYFDGKPPDMKQOELAKRYSKRDATKDLTEAVEVGD 120
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 KDALEKLSRTKYVTRQHNEDCKRLRLMGVNVPEAPSEAEACALCINDVFAVASED 180
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 K---MKQERRLVKYSKHENEAQKLTGLMGIPYLIAPTEAQAELAKKGYAASED 177
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 MDSLTFGAPRFLRLHMDPSSKRIPVMEFDYAKVLELELMDQIDLCILGCDYCSIR 240
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 MDCLTFGSPVLMRLHASEAKKLPIDFHLISRYLOELGINOQFVLDICILGSDYCESIR 237
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 GIGGOTAKLIRQHSISTELENKDRQIPEDMPYQEARLFEKPNVNTLDIELELWT 295
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 GIGAPVTAALKLITGHSIEKIEVFIESGESNNTKKIPEDMPYQEARLFDPEV-IDGNE 296
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 --LKWTPADEGLISFLVKGNGFNEDRYTKALEIKSAKNSSGRLSEFFKPTATTSAPLRKET 350
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 INLWSPPEKKELEIYLODDKRFSEERKSGISRLKGLASGIDGRLDGFQVVPKTEQ 356
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 -SAPLRKETSDKTSKAANKTK 373
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 LAAAKRAQENKKLK-NKNKYTK 379
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-08-823-516-139
; Sequence 139, Application US/08823516
; Patent No. 3994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichayev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
; TITLE OF INVENTION: Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Medgen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-823-516-139

Query Match 48.4%; Score 938; DB 2; Length 382;
Best Local Similarity 49.5%; Pred. No. 1,2e-84;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGIGLKRLLADNAPKMKQKRFESYFGRIKIVADSMSTIYOFLLIVGRTGMETLTNAGE 60
DB 1 MGIGLVAITSEHPSPAIRKSDIKSEFGKRVADSMSTLYOFLIVROODGGQLTNEGE 60
QY 61 VTSHLQGMFNRTIRLLGAIKPYVFDGKPPDMKQKOLAKRYSKRDDATKDLTEAVEYGD 120
DB 61 TTSLLMGFMFRTLLMDINGIKPCYVFDGKPPDLKSHETKRSRVRTEKKLAETTELE 120
QY 121 KDALEKSKRTYKVTROHNEDECKRLRLGNGVPEAPSEAECAALCTNDKVFVASED 180
DB 121 K---MKERRLIVYKSKHNEFAOKLLGLMGIPYITAPTEAKQCELAKKGVYAAASD 177
QY 181 MDSTLFAPRRLHMDPSKSLPVMEFYAKYLELELTMDQFDLTCLCGCDYDSIK 240
DB 178 MDVLCYTPPLRLHLLTFSEKKEPHIEIDTELVRLGLDITLIDFQDLTCLMGCCDCEISIR 237
QY 241 GIGGOTLKLIRHGSIESILENL-----NKDRYQIPEDMPYQEARLFEKPNVTLDIPE 295
DB 248 GVGVTALKLITKHGSEIEKIVEFESGESNNRKWIPEDMPYKQARMLFLDDEV-IDONE 296
QY 296 --LKWTPADDEGLISFLVKNNGENEDRVTKAIKIKSANNKSSQGRLESFEFTATT--- 350
DB 297 INLWSPPKKELELYLDCDKFSEERVKSGISRLKKGLKSGIQLDGFQVVPKTKEO 356
QY 351 -SAPLKRKETSDDTKSKAANKTKR 373
DB 357 LAAAKRAQENKRLNK-NKMKVTK 379

RESULT 13
US-08-757-653-176
Sequence 176, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-653-176

Query Match 33.4%; Score 648.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 4.7e-56;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KQKRESEYFGRIKIVADSMSTIYOFLLIVGRTGMETLTNAGEVTSHLQGMFNRTIRLLA 78
DB 12 KEIELENYGKRIADALNATYQFELSTROKDGTPLMDSKGRITSHLSGLFRTILMMEA 71
QY 79 GIKPYVFDGKPPDMKQKOLAKRYSKRDDATKDLTEAVEYGDATKDLTKSKRTYKVTROH 138
DB 72 GIKPYVFDGKPPDMKQKOLAKRYSKRDDATKDLTEAVEYGDATKDLTKSKRTYKVTROH 131
QY 139 NEDCKRLRLMGVPEAPSEAECAALCTNDKVFVASEDMDSLTFGAPRRLHMDP 198
DB 132 IDEAKLLELMGIPVQAPSEGEAAAYMAKGSVYASQDDVDSLTFGAPRRLHMDP 191
QY 199 SSKKIP-----VMEFDYAKVLELELTMDQFDLTCLCGCDYDSIK 244
DB 192 GKRLPKGKVVYKVEIPKELLILE---EVKLKELKTLRKLIELALIVGTQDYNPGIGIGL 247
QY 245 GTVALKILIRHGSIESILENLKD---RYQIPEDMPYQEARLFEKPNVTLDIPELKWTP 301
DB 248 KRALEIVRH-----SKDPLAKFQKOSQVDVLYAIKFEFLNLPVYTDNYNLVWROP 295
QY 302 DEEGLISFLVKNNGENEDRVTKAIKIKSANNKSSQGRLESFPR 345
DB 296 DEEGLIFLDCDEHDSERVYKNGLELTKAIKSGKOSTLESWEK 339

RESULT 14

US-08-823-516-79
: Sequence 79, Application US/08823516
: Patent No. 5994069
: GENERAL INFORMATION:
: APPLICANT: Hall, Jeff G.
: APPLICANT: Lyamichew, Victor I.
: APPLICANT: Mast, Andrea L.
: APPLICANT: Brow, Mary Ann D.
: TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
: TITLE OF INVENTION: Sequential Invasive Cleavages
: NUMBER OF SEQUENCES: 163
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Medlen & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States Of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/823,516
: FILING DATE: 24-MAR-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US97/01072
: FILING DATE: 21-JAN-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/759,038
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/758,314
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/756,386
: FILING DATE: 29-NOV-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/682,853
: FILING DATE: 12-JUL-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/599,491
: FILING DATE: 24-JAN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Ingolia, Diane E.
: REGISTRATION NUMBER: 40,027
: REFERENCE/DOCKET NUMBER: FORS-02736
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 79:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 340 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-823-516-79

Query Match 33.4%; Score 648.5; DB 2; Length 340;
Best Local Similarity 41.3%; Pred. No. 4.7e-56;
Matches 142; Conservative 62; Mismatches 107; Indels 33; Gaps 6;

QY 19 KEQFESYFGRIKIAVDASMSIYQFLIVGRTEMELTNEAGEVTSHTQGMFNRTIRLEA 78
Db 12 KEIETENLVGKTAIDALNATYQFSTIQKQGTPLMSKSRITSHLSGLFYRTINIMEA 71
QY 79 GIKRVYVFDKPPDMKKQELAKRYKRDATADLTENAEVGDKAIEKLSKRYKVTROH 138
Db 72 GIKRVYVFDKPPDMKKQELAKRYKRDATADLTENAEVGDKAIEKLSKRYKVTROH 131
QY 139 NEDCKRLILMGVYVFAESEAECALCINDKVFVAVASEMDSLTFGAPRFLRLHMDP 198

Db 132 IEDAKKLEIMGPIYQVAPSECEAOAAVMAGSVYASASQDYSLTFGAPRLVRLTIT 191
QY 199 SSKRIP-----VMEFDVAKVLELELTMOQFDITLCGCCD-CSTIGIG 244
Db 192 GKRLKQKNVYVYIKPELLILE---EYKELKLTREKLEIFALVGTIDYDPGIGIGL 247
QY 245 QYAKLIRQHGSIETILEMND---RYQIPEDMYPQAPRRLEPNVTLIDELKWTAP 301
Db 248 KKALEIVRH-----SKDPLAKFQKQSDVLYAIKEFFLMPVY-DNYNLVWRDP 295
QY 302 DEEGLISFLVKGNGFEDRVTKAIEKIKSAKKNSSQGRLESEFFK 345
Db 296 DEEGLISFLVKGNGFEDRVTKAIEKIKSAKKNSSQGRLESEFFK 339

RESULT 15
US-08-823-516-136
: Sequence 136, Application US/08823516
: Patent No. 5994069
: GENERAL INFORMATION:
: APPLICANT: Hall, Jeff G.
: APPLICANT: Lyamichew, Victor I.
: APPLICANT: Mast, Andrea L.
: APPLICANT: Brow, Mary Ann D.
: TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
: TITLE OF INVENTION: Sequential Invasive Cleavages
: NUMBER OF SEQUENCES: 163
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Medlen & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States Of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/823,516
: FILING DATE: 24-MAR-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US97/01072
: FILING DATE: 21-JAN-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/759,038
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/758,314
: FILING DATE: 02-DEC-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/756,386
: FILING DATE: 29-NOV-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/682,853
: FILING DATE: 12-JUL-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/599,491
: FILING DATE: 24-JAN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Ingolia, Diane E.
: REGISTRATION NUMBER: 40,027
: REFERENCE/DOCKET NUMBER: FORS-02736
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 136:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 340 amino acids
: TYPE: amino acid

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:26:59 ; Search time 21 Seconds

(without alignments)
1734.184 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939

Sequence: 1 MGIKGLTKLADNAPKAMKE.....SDKTSKAANKTKAGGKK 379

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR-71:*
2: PIR1:*
3: PIR2:*
4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1313.5	67.7	362	2 T01198	endonuclease homol
2	1032	53.2	380	2 A56531	DNA structure-spec
3	1023.5	52.8	380	2 A56054	DNA repair endonuc
4	1014	52.3	378	2 A53730	flap endonuclease
5	998.5	51.5	385	2 T13692	hypothetical prote
6	938	48.4	382	2 S22267	hypothetical prote
7	648.5	33.4	340	2 T46893	endo/exonuclease f
8	636	32.8	343	2 E75117	DNA repair protein
9	632.5	32.6	343	2 A71015	probable 5' nuclea
10	560.5	28.9	326	2 C64480	DNA repair protein
11	542	28.0	328	2 C69085	DNA repair protein
12	538.5	27.8	336	2 H69282	DNA repair protein
13	514.5	26.5	401	2 H72765	probable flap endo
14	506	26.1	302	2 F90158	DNA repair endo/ex
15	369	19.0	327	2 F84290	DNA repair protein
16	336	17.3	1516	2 E71619	RAD2 endonuclease
17	267.5	13.8	317	2 F86419	probable exonuclea
18	260.5	13.4	571	2 T43288	probable exodeoxyr
19	245.5	12.7	496	2 S58300	probable excision
20	235	12.1	457	2 F86150	F22M8.2 protein -
21	232.5	12.0	1031	2 A29839	RAD2 protein - yea
22	220.5	11.4	1186	2 S35993	DNA repair protein
23	220	11.3	1170	2 A57650	repair protein XPG
24	218	11.2	702	2 S62169	DH51 protein - yea
25	206	10.6	1113	2 S30301	excision repair pr
26	204.5	10.5	430	2 S61118	DIN7 protein - yea
27	204.5	10.5	1112	2 T40382	dna repair protein
28	202	10.4	1185	2 T58009	gene EXC5 protein
29	200	10.3	1196	2 S35994	DNA repair protein

30	197	10.2	1287	2 T22235	hypothetical prote
31	179.5	9.3	829	2 T32744	hypothetical prote
32	176.5	9.1	986	2 S77056	DNA-directed DNA p
33	173	8.9	832	2 A33530	DNA-directed DNA p
34	172.5	8.9	434	2 T34350	hypothetical prote
35	171	8.8	831	2 JX0359	DNA-directed DNA p
36	166	8.6	289	2 D70440	DNA polymerase I
37	162.5	8.4	871	2 E97035	DNA polymerase I
38	158.5	8.2	290	2 A11309	5'-3' exonuclease
39	158.5	8.2	492	2 C86316	protein T10022.7 l
40	157.5	8.1	831	2 S26675	DNA-directed DNA p
41	156.5	8.1	877	2 A32949	DNA-directed DNA p
42	156.5	8.1	877	2 H97875	DNA-directed DNA p
43	154.5	8.0	877	2 E95003	DNA polymerase I
44	151.5	7.8	337	2 T49286	hypothetical prote
45	151.5	7.8	759	2 S50544	hypothetical prote

ALIGNMENTS

RESULT 1

T01198

endonuclease homolog F21E10.3 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999

C:Accession: T01198

R:Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.

submitted to the EMBL Data Library, April 1998

A:Description: The sequence of A. thaliana F21E10.

A:Reference number: Z14258

A:Accession: T01198

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-362 <DAV>

A:Cross-References: EMBL:AF058914; NID:93047074; PTD:93047083; PIDN:AA033596.1; GSPDB

A:Experimental source: cultivar Columbia

C:Gene(s):

A:Gene: ATSP:F21E10.3

A:Map position: 5

A:Introns: 4/3: 51/2: 71/1: 96/2: 116/2: 134/2: 150/3: 166/3: 188/3: 205/3: 245/3: 27

C:Superfamily: conserved hypothetical protein YKL113c

C:Keywords: DNA repair; endonuclease

Query Match	67.7%	Score 1313.5; DB 2; Length 362;
Best Local Similarity	72.1%	Pred. No. 1.6e-80;
Matches 259; Conservative 30; Mismatches 33; Indels 37; Gaps 3;		
QY 1 MGIKGLTKLADNAPKAMKEKFESYFGFKIAVDASMSIYQFLIVGRTGMYLTNAGE 60		
1 MGIKGLTKLADNAPKAMKEKFESYFGFKIAVDASMSIYQFLIVGRTGMYLTNAGE 60		
DB 1 MGIKGLTKLADNAPKAMKEKFESYFGFKIAVDASMSIYQFLIVGRTGMYLTNAGE 60		
QY 61 VTS-----HLGEMFRITRLLEAGIKPYV 86		
61 VTS-----HLGEMFRITRLLEAGIKPYV 86		
DB 60 KCSLMKLVSLWIFPLDGLCLYFVVVYIMMNRHLQMFNRTIRLEAGIKPYV 119		
QY 87 DGKPPMKKQELAKRYSKRDATKDLTEAVEVGDALEKLSKRTVAVTROHNEDEKRL 146		
87 DGKPPMKKQELAKRYSKRDATKDLTEAVEVGDALEKLSKRTVAVTROHNEDEKRL 146		
DB 120 DGKPPMKKQELAKRYSKRDATKDLTEAVEVGDALEKLSKRTVAVTROHNEDEKRL 146		
QY 147 RLMGVVVAAPSAPEACALCINDKRVAVASEMDSTTEAPFRLHMLDPSSKRTIPV 206		
147 RLMGVVVAAPSAPEACALCINDKRVAVASEMDSTTEAPFRLHMLDPSSKRTIPV 206		
DB 180 RLMGVVVAAPSAPEACALCINDKRVAVASEMDSTTEAPFRLHMLDPSSKRTIPV 206		
QY 207 EPDVAVLELELTMDQFIDLCILGCDYCDISIKIGGQALKILRHGSIESTILENLK 266		
207 EPDVAVLELELTMDQFIDLCILGCDYCDISIKIGGQALKILRHGSIESTILENLK 266		
DB 240 EPDVAVLELELTMDQFIDLCILGCDYCDISIKIGGQALKILRHGSIESTILENLK 266		
QY 267 DRGQIEDMPYQEARLFEKPVVTLIDPEL--KWPAPREELSELKDGNGFNDPRTK 323		
267 DRGQIEDMPYQEARLFEKPVVTLIDPEL--KWPAPREELSELKDGNGFNDPRTK 323		
DB 300 ERQIQEEMPIEAKRLEKEPVITDEQDLIKTISPEDEGVQFLVNEGFNIDRYTK 358		

R:Harrington, J.J.; Lieber, M.R.

Genes Dev. 8, 1344-1355, 1994

A:Title: Functional domains within FEN-1 and RAD2 define a family of structure-specific

A:Reference number: A53730; MUID:95011546

A:Accession: A53730

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-378 <HAR>

A:Cross-references: GB:U26320; NID:g499648; PIDN:AAC37664.1; PID:g499649

C:Superfamily: conserved hypothetical protein YKL113c

C:Keywords: DNA repair; endonuclease

Query Match 52.3%; Score 1014; DB 2; Length 378;

Best Local Similarity 53.2%; Pred. No. 1.8e-60;

Matches 202; Conservative 69; Mismatches 101; Indels 8; Gaps 6;

QY 1 MGIGLTKLADNAPKAEQKESYFGKRIAVDASMSIYOFLLVVGRTGEMTLTNEAGE 60
 DB 1 MGIGLAKLADVAPSAIRENDIKSYFGKRYAIDASMSIYOFLLVAV-RGGDVLYQDEGE 59
 QY 61 VTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKROELAKRYSKRDATKDLTEAVEVGD 120
 DB 60 TTS-LMGMYFRTIR-MENGIRPVYFDGKPPQLKSGELAKRSEERRAEKQLQAQAEAGM 117
 QY 121 KAIEKLSKRTVYTRHONEDCKRLRLMGVYVPAEAPSEABCAALCINDKYFAVAASED 180
 DB 118 EEEVEKFTRLVYTKRKHNDCKHLISLMIPIYLDAPSEABCAALAKAKKYAATATED 177
 QY 181 MSLTFGAPRFLRLHMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILCGDCYCSIK 240
 DB 178 MCLTFGSPVLMRLHJLASEAKRLPQIEFHLISRVLOELGNOQFVDCILGSDYCESIR 237
 QY 241 GIGGQATLKLIRHGSIESILENKDRYOIPEDMPYQEARLJFKEPNVT-LDIPELKMT 299
 DB 238 GIGAKRAVDLLOKHSISEIYRRLDPKSYPEEWMLKREKQOLFEEVVDPSVELKMS 297
 QY 300 APDEGLISFLVKNDFNEDRYTKAIEKIKSAKNSOGRLSEFFKPTATTSAPLKRKET 359
 DB 298 EPNEELVAFKCEKQFSEERIRSGVKLSRSGSTQGRDLDFKVTGLSL-ARRKET 356
 QY 360 SDKTSKAANKKTKAGGKK 379
 DB 357 E---PKGPAKKAKTKGAGK 373

RESULT 5

hypothetical protein EG0003.3 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C:Accession: T13692

R:Murphy, L.; Harris, D.; Barrell, B.

submitted to the EMBL Data Library, November 1998

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17699

A:Accession: T13692

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-385 <MDR>

A:Cross-references: EMBL:AL031863; PIDN:CAA21320.1

C:Genetics:

A:Cross-references: FlyBase:FBgn0025833

A:Introns: 28/1

A>Note: EG:EG0003.3

C:Superfamily: conserved hypothetical protein YKL113c

Query Match 51.5%; Score 998.5; DB 2; Length 385;

Best Local Similarity 52.9%; Pred. No. 2e-59;

Matches 202; Conservative 65; Mismatches 104; Indels 11; Gaps 5;

QY 1 MGIGLTKLADNAPKAEQKESYFGKRIAVDASMSIYOFLLVVGRTGEMTLTNEAGE 60
 DB 1 MGIGLTKLADLQALRESEMKHFEGRKVALDASMCYQFLVAVSEAG-LATVIND 59

QY 61 VTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKROELAKRYSKRDATKDLTEAVEVGD 120
 DB 60 PTHLMGMFYRTIRLLDNGIKRPVYFDGKPPDLKSGELAKRREERAEKALKMATDAG 119
 QY 121 KDAIEKLSKRTVYTRHONEDCKRLRLMGVYVPAEAPSEABCAALCINDKYFAVAASED 180
 DB 120 DAGEKLNRLVYTKHAKKELLTLMGVYVPAEAPSEABCAALAKAGRYATATED 179
 QY 181 MSLTFGAPRFLRLHMDPSSKIPVMEFDVAKVLELELTMDQFIDLCILCGDCYCSIK 240
 DB 180 MDALTFGSTKLLLYLYTSEARKMPVEFSYDKLGLAINNREFIDLCILCGDCYCSIK 239
 QY 241 GIGGQATLKLIRHGSIESILENKDRYOIPEDMPYQEARLJFKEPNVT-LDIPELKMT 299
 DB 240 GIGKRAIELINRYRDIETLLDMLDSSKTYPEENNYKARLFLFEVVAADSIDLKW 299
 QY 300 APDEGLISFLVKNDFNEDRYTKAIEKIKSAKNSOGRLSEFFK---PTATTSAPL 354
 DB 300 EPDEGLVAKFLCDROFNEERYVNGAKLMSKQAOYQVRLDSEFKTLPSTPNATNA-- 357
 QY 355 KRKETSCKAANKKTRAG 376
 DB 358 KRK--AEAKKSAANKKATSG 377

RESULT 6

hypothetical protein YKL113c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YKL510; RAD2 protein homolog

C:Species: Saccharomyces cerevisiae

C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 24-Sep-1999

C:Accession: S22267; S37941

R:Jacquier, A.; Legrain, P.; Dujon, B.

Yeast 8, 121-132, 1992

A:Title: Sequence of a 10.7 kb segment of yeast chromosome XI identifies the APN1 and

S.

A:Reference number: S22267; MUID:92221689

A:Accession: S22267

A:Molecule type: DNA

A:Residues: 1-382 <JAC>

A:Cross-references: GB:S93804; NID:g248391; PIDN:AAB21998.1; PID:g248392

A:Experimental source: strain S288C

R:Jacquier, A.; Legrain, P.; Colletaux, L.; Richard, G.F.; Thierry, A.; Dujon, B.

submitted to the Protein Sequence database, March 1994

A:Reference number: S37938

A:Accession: S37941

A:Molecule type: DNA

A:Residues: 1-382 <JAC>

A:Cross-references: EMBL:Z28113; NID:g486189; PIDN:CAA81953.1; PID:g486190; MIPS:YKL1

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SCD:RAD27

A:Cross-references: SGD:S0001596; MIPS:YKL113c

A:Map position: 11L

C:Superfamily: conserved hypothetical protein YKL113c

C:Keywords: nucleus

Query Match 48.4%; Score 938; DB 2; Length 382;

Best Local Similarity 49.5%; Pred. No. 2.2e-55;

Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;

QY 1 MGIGLTKLADNAPKAEQKESYFGKRIAVDASMSIYOFLLVVGRTGEMTLTNEAGE 60
 DB 1 MGIGLNAVITSEHPSPAIRSDIKSFGRKVAIDASMSIYQFLVAVRQDGGQLTNEAGE 60
 QY 61 VTSHLQGMFNRTIRLEAGIRPVYFDGKPPDMKROELAKRYSKRDATKDLTEAVEVGD 120
 DB 61 TTSMLGMFYRTIRLMDINGIKRPYVFDGKPPDLKSGELAKRSEERRAEKALATATELE 120
 QY 121 KDAIEKLSKRTVYTRHONEDCKRLRLMGVYVPAEAPSEABCAALCINDKYFAVAASED 180
 DB 121 K---MKGRRLVAVSKENHNEAOKLLGLMGIPYITAPTEAEOCAELAKKAKRYAASSED 177

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-327 <STO>

A:Cross-references: GB:AE004437; NID:g10580870; PIDN:AMG19690.1; GSPDB:GN00138

C:Genetics:

A:Gene: rad2

C:Superfamily: conserved hypothetical protein YK113c

Query Match 19.0%; Score 369; DB 2; Length 327;

Best Local Similarity 31.8%; Pred. No. 2e-17;

Matches 106; Conservative 60; Mismatches 139; Indels 28; Gaps 9;

```
Oy 17 AMKQKFESEYFGRIKIAVDASMSIYOF-IYVGRGMEITLNEAGEVTSILOGMFNRTIRL 75
Db 11 AIEETPFADLEGSVAVADAHNMILYKYLITTVQMTGADVTTSDGTEVANLVGAVGCLPKF 70
Oy 76 LEAGIKRYVYFDGKFPDMKKQELAKRISKRDATKDLTEAVEYGDKAIEKSKRTVAVT 135
Db 71 FEHGILTFVFWDDGGVTELKDDDEIADRREQREYEEQLDDAREGDAAEAARLIDARTORLT 130
Oy 136 RQHNEDCKRLRLKMGVPEVPEAEAECAALC-INDKYFAVASEDMDSLTFGAPRFLRH 194
Db 131 PTIHETTFELFDLDIPQVEAPAEGEQAAYMTRTDDAVDAGSDDYDCLLGSPVTLRQ 190
Oy 195 LMDPSKKIPWFERPDVAKVLELEELTMDQFIDLCITCGGDCDSYKGIQGTALKLIRQH 254
Db 191 L--TSSGHPELMDPDA--TLAEHDLWEQLVDVGLIGTDFNPIDGFGFTTALDAIGBH 246
Oy 255 GSIESITL---ENL-NKDRYQIPEDWPEQEARRLFEKPEPVTLDI---PELKWTPADEEGL 306
Db 247 GDLMDVTLAEGEHVAHGDR-----IRELFLNPVTDVYIDPV---SPALDAA 292
Oy 307 ISFLVKONGFENEDRVTKAIEKIKSAKKKSSQGR 339
Db 293 RAFVTDMEWVDADAVARGFERIDAAAQGTGDR 325
```

Search completed: November 5, 2002, 16:29:44

Job time : 24 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:25:09 : Search time 15 seconds

(without alignments)
978.314 Million cell updates/sec

Title: US-09-805-311-2

Sequence: 1 MGIGLTKLADNAPKAKE.....SDKTSKAAANKTKRAGKKK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1032	53.2	380	1	FEN1_HUMAN
2	1023.5	52.8	380	1	RAD2_SCHPO
3	1014	52.3	378	1	FEN1_MOUSE
4	998	48.4	382	1	RA27_YEAST
5	560.5	28.9	326	1	FEN1_METUA
6	260.5	13.4	571	1	EXO1_SCHPO
7	245.5	12.7	496	1	YA31_SCHPO
8	232.5	12.0	1031	1	RAD2_YEAST
9	220.5	11.4	1186	1	XPG_HUMAN
10	220	11.3	1170	1	XPG_MOUSE
11	218	11.2	702	1	EXO1_YEAST
12	204.5	10.3	430	1	DIN7_YEAST
13	204.5	10.5	1112	1	RA13_SCHPO
14	200	10.3	1196	1	XPG_XENLA
15	184.5	9.5	834	1	DPOL_THECA
16	176.5	9.1	986	1	DPOL_SYNY3
17	173	8.9	832	1	DPOL_THEAO
18	172.5	8.9	834	1	DPOL_THEHT
19	166	8.6	289	1	EX53_AOUAE
20	157.5	8.1	831	1	DPOL_THEFL
21	156.5	8.1	833	1	DPOL_THEFT
22	154.5	8.0	877	1	DPOL_STPRN
23	153.5	7.9	877	1	DPOL_LACIC
24	151.5	7.8	759	1	YEN1_YEAST
25	149	7.7	930	1	DPOL_HAEIN
26	144	7.4	877	1	DPOL_BACCA
27	142.5	7.3	877	1	DPOL_LACIA
28	138	7.1	921	1	DPOL_RICHE
29	135.5	7.0	880	1	DPOL_BACSU
30	134.5	6.9	956	1	DPOL_DEIRA
31	134	6.9	867	1	DPOL_RICPR
32	134	6.9	922	1	DPOL_RICFE
33	123	6.3	876	1	DPOL_BACST

34	118.5	6.1	850	1	DPOL_ANATH	059156 anaerocellu
35	113.5	5.9	942	1	DPOL_CHIAU	008307 chloroflexu
36	112	5.8	317	1	ACCO_PEA	P31239 pisum sativ
37	112	5.8	984	1	SECA_AOUAE	067718 aquilex aeo
38	109.5	5.6	891	1	DPOL_HELPU	P56105 heliobacte
39	109.5	5.6	1016	1	DPOL_RHILE	Q98192 rhizobium 1
40	109.5	5.6	1650	1	BP28_CAEEL	Q23495 caenorhabd1
41	109	5.6	886	1	RA50_ARCEU	029230 archaeglob
42	109	5.6	897	1	DPOL_HELPU	092199 heliobacte
43	108.5	5.6	444	1	TIG_CAMTE	Q46108 campylobact
44	108.5	5.6	928	1	DPOL_ECOLI	P00582 escherichia
45	108.5	5.6	1509	1	MISN_CACCA	P05659 acanthameb

ALIGNMENTS

```

RESULT 1
ID      FEN1_HUMAN      STANDARD:      PRT:      380 AA.
AC      P39748:
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      FLAP endonuclease-1 (Maturation factor 1) (MFI).
GN      FEN1.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NC      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94277093; PubMed=8007985;
RA      Murray J.M., Tavassoli M., Al-Harithy R., Sheldrick K.S.,
RA      Lehmann A.R., Carr A.M., Watts F.Z.;
RT      "Structural and functional conservation of the human homolog of the
RT      Schizosaccharomyces pombe rad2 gene, which is required for chromosome
RT      segregation and recovery from DNA damage.";
RL      Mol. Cell. Biol. 14:4878-4888(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      TISSUE=Lymphocytes;
RA      Lieber M.R., Hsieh C.L., Harrington J.J.;
RL      Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Lamerdin J.E., McCreedy P.M., Coleman M., Skowronski E., Adamson A.W.,
RA      Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA      Phan H., Velasco N., Do L., Regala W., Terry A., Gaines J.,
RA      Dangnan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA      Liu S., Attix C., Andrieux T., Frankheim M., Amico-Keller G.,
RA      Koefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA      Krommiller B., Ariello A., Montgomery M., Ow D., Nolan M., Truong S.,
RA      Kobayashi A., Olsen A.S.;
RA      Carrano A.V.;
RL      Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RX      TISSUE=Lung;
RA      Strausberg R.;
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      PARTIAL SEQUENCE.
RX      MEDLINE=95050647; PubMed=7961795;
RA      Robins P., Pappin D.J., Wood R.D., Lindahl T.;
RL      "Structural and functional homology between mammalian Dnase IV and
RL      the 5'-nuclease domain of Escherichia coli DNA polymerase I.";
RL      J. Biol. Chem. 269:28535-28538(1994).
CC      -I- FUNCTION: ENDONUCLEASE THAT CLEAVE 5'FLAP STRUCTURE AND FAILS
CC      TO CLEAVE OTHER DNA STRUCTURES, INCLUDING 3'FLAPS AND SINGLE
CC      STRANDED DNA (BY SIMILARITY).
CC      -I- SUBCELLULAR LOCATION: Nuclear (Probable).
CC      -I- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
  
```

```

CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X76771; CA54166.1; -
DR      EMBL: L37374; AAA91331.1; -
DR      EMBL: AC004770; AAC2394.1; -
DR      EMBL: BC000323; AAA00323.1; -
DR      HSSP: Q58839; 1A76.
DR      MIM: 600393; -
DR      InterPro: IPR002421; 5_3_exonuclease.
DR      InterPro: IPR000513; Exo_N.T.
DR      InterPro: IPR003584; HNH_2.
DR      InterPro: IPR001532; XPG_1.
DR      Pfam: PF01367; 5_3_exonuclease; 1.
DR      Pfam: PF00867; XPG_1; 1.
DR      Pfam: PF00752; XPG_N; 1.
DR      PRINTS: PR00853; XPGRADSUPER.
DR      SMART: SM00279; HNH2; 1.
DR      SMART: SM00484; XPG1; 1.
DR      SMART: SM00485; XPGN; 1.
DR      PROSITE: PS00841; XPG_1; 1.
DR      PROSITE: PS00842; XPG_2; 1.
DR      Hydrolase; Nuclease; Endonuclease; Nuclear protein.
KT      DOMAIN 1 104
FT      DOMAIN 122 253
SQ      SEQUENCE 380 AA; 42593 MW; 5154PF6E5792C5 CRC64;

Query Match          53.2%; Score 1032; DB 1; Length 380;
Best Local Similarity 53.8%; Pred. No. 3,1e-61;
Matches 206; Conservative 69; Mismatches 98; Indels 10; Gaps 6;

QY      1 MGIGKLTKLADNAPKAMKEQKFESEYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
DB      1 MGIGGLAKLTIADVAPSAIRENDIKSYFKRKVAIDASMSIYQFLIIV-RGGGVLIQNEGE 59

QY      61 VTSHLOGMFKRTIRLEAGIKPRYYVDGKPRPMKKOELAKRKSKRDKDTKDLTEAVEGCD 120
DB      60 TTSHLMGMFKRTIRLMENGGIKPRYYVDGKPRPOLSKSELAKRSEKRAEAKLOQQAAGA 119

QY      121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVPEVPEAPSEAEACALCINDKVFVAVASED 180
DB      120 EQEVEKFTKRLVVTQHNDECKHLLSLMGIRPLDAPSEAEACALVAKGVVAATED 179

QY      181 MDSLTFGARPLRLHMDPSSKKIPVWEFDVAKVLELELTMDQFLDLCILGCGDYCSIK 240
DB      180 MDSLTFGSPVLRHRLVASEKKLPIQEFHLISRLDGLNOGFYDLCLILGSDCESIR 239

QY      241 GIGGOTAKLTIROHGSIESILENLNKRQYQIPEDWPEQARLFKEPNVTTD--IPELKW 298
DB      240 GIGKRAVVDLIQNHKSIEELVRLDNPKNKYRPPENMLHKAHOLFLEPEV-LDPEVELKW 298

QY      299 TAPDEGLIFLVKDNQFNDRYTKAIEKIKSAKNNSSQGRLESPFKPATTSAPLAKKE 358
DB      299 SEPHEELIFMCGEKQFSEERIRSGVKRLSKRSGSTOGRLDPEFKYGLSS--AKRKE 357

QY      359 TSDK--TSKAANKTKAGGKK 379
DB      358 PERKGTAKKA---KTGAAGKFK 377

```

```

DT      01-MAR-2002 (rel. 41, Last annotation update)
DE      DNA repair protein rad2.
GN      RAD2 OR SPAC36.06C.
OS      Schizosaccharomyces pombe (Fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomycetes.
OX      NCBI_TaxID=4896;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9427093; PubMed=8007985;
RA      Murray J.M., Tavassoli M., Al-Harithy R., Sheldrick K.S.,
RA      Lehmann A.R., Carr A.M., Watts F.Z.;
RT      Structural and functional conservation of the human homolog of the
RT      Schizosaccharomycetes pombe rad2 gene, which is required for chromosome
RT      segregation and recovery from DNA damage.;
RL      Mol. Cell. Biol. 14:4878-4888(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=972;
RA      Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL      Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC      FUNCTION: REQUIRED FOR THE FIDELITY OF CHROMOSOME SEPARATION AT
CC      MITOSIS AND WHICH IS ALSO INVOLVED IN THE RESPONSE TO DNA DAMAGE.
CC      -1- SUBCELLULAR LOCATION: Nuclear (probable).
CC      -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENONUCLEASE FAMILY. FEN1
CC      SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X77041; CAB36991.1; -
DR      EMBL: Z99167; CAB16282.1; -
DR      HSSP: Q58839; 1A76.
DR      InterPro: IPR000513; Exo_N.1.
DR      InterPro: IPR003584; HNH_2.
DR      InterPro: IPR001532; XPG_1.
DR      Pfam: PF00867; XPG_1; 1.
DR      Pfam: PF00752; XPG_N; 1.
DR      PRINTS: PR00853; XPGRADSUPER.
DR      SMART: SM00279; HNH2; 1.
DR      SMART: SM00484; XPG1; 1.
DR      SMART: SM00485; XPGN; 1.
DR      PROSITE: PS00841; XPG_1; 1.
DR      PROSITE: PS00842; XPG_2; 1.
DR      DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
FT      DOMAIN 1 105
FT      DOMAIN 123 254
SQ      SEQUENCE 380 AA; 42866 MW; 36040659DB934CBF CRC64;

Query Match          52.8%; Score 1023.5; DB 1; Length 380;
Best Local Similarity 50.7%; Pred. No. 1,1e-60;
Matches 194; Conservative 73; Mismatches 107; Indels 9; Gaps 3;

QY      1 MGIGKLTKLADNAPKAMKEQKFESEYFGKRIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
DB      1 MGIGLAQVLSERHAPASVKNRDKIKVFGRKVAIDASMSIYQFLIIVRSODQQLMNEGE 60

QY      61 VTSHLOGMFKRTIRLEAGIKPRYYVDGKPRPMKKOELAKRKSKRDKDTKDLTEAVEGCD 120
DB      61 TTSHLMGMFKRTIRLDVNGIKPCFVFDGKPRPLKSGELAKRYARQKAREQDEETKEVGT 120

QY      121 KDAIEKLSKRTYKVTROHNEDCKRLRLMGVPEVPEAPSEAEACALCINDKVFVAVASED 180
DB      121 AAVVDFAKRTYVTRQHNDEAKRLLELMGIRPVNAPCEAEACALARSQVYAAASED 180

QY      181 MDSLTFGARPLRLHMDPSSKKIPVWEFDVAKVLELELTMDQFLDLCILGCGDYCSIK 240
DB      181 MDSLTFGARPLRLHMDPSSKKIPVWEFDVAKVLELELTMDQFLDLCILGCGDYCSIK 240

```



```

CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S93804; AAB21998.1; -
CC DR EMBL: Z28113; CAAB1953.1; -
CC DR PIR: S22267; S22267.
CC DR HSSP: O58839; 1A76.
CC DR SGD: S0001596; RAD27.
CC DR InterPro: IPR002421; 5_3_exonuclease.
CC DR InterPro: IPR000513; EXO_N1.
CC DR InterPro: IPR003584; HHN_2.
CC DR InterPro: IPR001532; XPG_1.
CC DR Pfam: PF01367; 5_3_exonuclease; 1.
CC DR Pfam: PF00867; XPG_1; 1.
CC DR Pfam: PF00752; XPG_N; 1.
CC DR PRINTS: PRO0853; XGRADSUPER.
CC DR SMART: SM00279; HHN2; 1.
CC DR SMART: SM00484; XPG1; 1.
CC DR SMART: SM00485; XPGN; 1.
CC DR PROSITE: PS00841; XPG_1; 1.
CC DR PROSITE: PS00842; XPG_2; 1.
CC DR DNA_repair: Hydrolase; Nuclease; Endonuclease; Nuclear protein.
CC FT DOMAIN 1 105 N-DOMAIN.
CC FT SEQUENCE 382 AA; 43279 MW; 1F54B08720121C8C CRC64;
SQ
Query Match 48.4%; Score 938; DB 1; Length 382;
Best Local Similarity 49.5%; Pred. No. 5e-55;
Matches 190; Conservative 69; Mismatches 109; Indels 16; Gaps 6;
OY 1 MGIGITLLADNAPKAKKEQFESYGRKTAVDASITVFLVGRGTMETLTNAGE 60
DB 1 MGIGKLNATISEHVPATRKSDIKSFGRKVAIDNSLVIYFLAVKQDGGGLTNAGE 60
OY 61 VTSHLOGMFNTIRLLLEAGIRPVYFDGKPPDMKKOELAKRYKRDATKDLFEAVEVG 120
DB 61 TSHLMGMFYRLRMIDGICPCYVFDGKPPDLKSHLTKRSSRVEKELAEATTELE 120
OY 121 KDAIEKLSKRYKYTRQINEDCKRLRLMGVPVYEAPEAEACALCINDKVAVAASED 180
DB 121 K---KKOERRLVKYSKEHNEEAOKLLGMLPIYIIAPEAEAOCAELAKKGVYAAASED 177
OY 121 K---KKOERRLVKYSKEHNEEAOKLLGMLPIYIIAPEAEAOCAELAKKGVYAAASED 177
OY 181 MSLTFGAPRFLRHLMDPSSKKIPWEDVAKVLEELTMDQFIDLCILGCDVCSIK 240
DB 178 MDTLCYRPPFLRLHLPSEAKKEPHIEDTLVLRGLDITLIEQFDLCIMGCGYCESTR 237
OY 241 GIGGOTALKLIHQHSISILENT-----NKDRQIPEDWPYQEARRLFEPNVTLDIPE 295
DB 241 GIGGOTALKLIHQHSISILENT-----NKDRQIPEDWPYQEARRLFEPNVTLDIPE 295
OY 238 GVGPTALAKLITGHSIEKIVFIESGESNNTKWKIPEDWPKQARMLFDLPEV-IDENE 296
DB 238 GVGPTALAKLITGHSIEKIVFIESGESNNTKWKIPEDWPKQARMLFDLPEV-IDENE 296
OY 296 --LKTATPEDEGLISFLVKNDFNEDRYTKATEIKSAKNNSSORLSEFPKPAAT--- 350
DB 297 INLKSPPEKEKLELYLDCDDKFESEKSGISRLKGLKSGISGRIDGFFOVVPKTKEQ 356
OY 351 -SAPLKRKETSQTSKAANKRKT 373
DB 357 LAAAKRAQENKKLNK-NKNKYTK 379

```

```

DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Flap endonuclease.
GN FEN OR M0144.
OS Methanococcus jannaschii.
CC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
CC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A. 2661 / ATCC 43067;
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Adams M.D., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Feldman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Huft M.A., Kalne B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RN Science 273:1058-1073(1996).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=98440432; PubMed=9765572;
RA Rao H.G., Rosenfeld A., Wetmur J.G.;
RT "Methanococcus jannaschii flap endonuclease: expression, purification,
RT and substrate requirements";
RN J. Bacteriol. 180:5406-5412(1998).
RN [3]
RP CHARACTERIZATION, AND X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=99200693; PubMed=10102570;
RA Bae K.W., Bae K.W., Cho C.S., Hwang K.Y., Kim H.-R., Sung H.-C.,
RA Cho Y.;
RT "Expression, purification, characterization and crystallization of
RT flap endonuclease-1 from Methanococcus jannaschii.";
RN Mol. Cells 9:45-48(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=98363215; PubMed=9699635;
RA Hwang K.Y., Bae K., Kim H.-Y., Cho Y.;
RT "The crystal structure of flap endonuclease-1 from Methanococcus
RT jannaschii.";
RN Nat. Struct. Biol. 5:707-713(1998).
RN [5]
RP FUNCTION: Has 5' endo-/exonuclease and 5' pseudo-Y-endonuclease
CC activities. Cleaves the junction between single and double-
CC stranded regions of flap DNA.
CC -1- COFACTOR: MAGNESIUM.
CC -1- MISCELLANEOUS: ITS OPTIMUM PH IS 6-7.
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67585; AAB99454.1; -
CC DR PDB: 1A76; 03-AUG-99.
CC DR PDB: 1A77; 03-AUG-99.
CC DR TIGR: M01444; -
CC DR InterPro: IPR000513; EXO_N1.
CC DR InterPro: IPR003584; HHN_2.
CC DR InterPro: IPR001532; XPG_1.
CC DR Pfam: PF00867; XPG_1; 1.
CC DR Pfam: PF00752; XPG_N; 1.
CC DR PRINTS: PRO0853; XGRADSUPER.
CC DR SMART: SM00279; HHN2; 1.
CC DR SMART: SM00484; XPG1; 1.
CC DR SMART: SM00485; XPGN; 1.

```



```

RP SEQUENCE FROM N.A.
RC STRAIN-972:
RA Devlin K., Churche C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z50728; CAA90586.1; -
DR InterPro: IPR000513; Exo_N.1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_1.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
KW Hypothetical protein; Hydrolase; Nuclease; Endonuclease.
SQ SEQUENCE 496 AA; 57763 MW; F2A16D26846F24C3 CRC64;

Query Match 12.7%; Score 245.5; DB 1; Length 496;
Best Local Similarity 30.2%; Pred. No. 3.3e-09;
Matches 62; Conservative 47; Mismatches 77; Indels 19; Gaps 5;

OY 111 DLTAEEVGDKADEIKSKRTAVYVTOHNEDECKRLRLMGVPEAP--SEAEKCALC 168
DB 231 ELQRAIKLKYTE-LDKLERLYRSPONIFELFELIKLIGIPASTSPGVAEAFSAIS 309
OY 169 INDVFAVASEDMDSLTFGAPRFLRLIMPSSKIPVMEFVAKYLELELTMOQFDLC 228
DB 310 ONNLAYAVATQDFVLLGSSMSINFLDNDNFHLPLQIMPRKIAQNLNLTFFGPDYC 369
OY 229 ILCCGDCDSIKGIGGTALKLRHGSLESTLENIN-KDRYQIPEDW--PYQEARLTK 285
DB 370 LMCGTDFTSRPTKIGPYRALKLIRYGNAPDYKALNVEKILPTDIKKFLTKAKLFT 429
OY 286 EPNVTLDIPELKWAPDEGLISFL 310
DB 430 D-----LPSNNELFSEFI 441

RESULT 8
RAD2_YEAST STANDARD; PRT; 1031 AA.
AC P07276;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA repair protein RAD2.
GN RAD2 OR YGR258C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86083164; Pubmed=3000874;
RA Nicolet C.M., Chenevert J.M., Friedberg E.C.;
RT "The RAD2 gene of Saccharomyces cerevisiae: nucleotide sequence and
RT transcript mapping.";
RL Gene 36:225-234(1985).
RN (2)
RP REVISIONS. SEQUENCE FROM N.A.
RA MEDLINE=86223812; Pubmed=3011752;
RA Madura K., Prakash S.;

```

```

RT "Nucleotide sequence, transcript mapping, and regulation of the RAD2
RT gene of Saccharomyces cerevisiae.";
RL J. Bacteriol. 166:914-923(1986).
RN [3]
RP SEQUENCE FROM N.A.
RA Panzeri L., Agostoni Carbone M.L., Melchiorre P., Plevani P.,
RA Martegani E., Vancori M., Carignani G., Clemente M.L., Frontali L.,
RA Fabiani L., Marconi A., Ruzzi M., Sallola M.;
RN Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-51 FROM N.A.
RC STRAIN=5288C / FY1679;
RX MEDLINE=97245300; Pubmed=9090059;
RA Clemente M.L., Sartori G., Cardazzo B., Carignani G.;
RT "Analysis of an 11.6 kb region from the right arm of chromosome VII of
RT Saccharomyces cerevisiae between the RAD2 and the MSL1 genes reveals
RL the presence of three new genes.";
RN Yeast 13:287-290(1997).
RN [5]
RP FUNCTION.
RX MEDLINE=94067324; Pubmed=8247134;
RA Habraken Y., Sung P., Prakash S.;
RT "Yeast excision repair gene RAD2 encodes a single-stranded DNA
RT endonuclease.";
RL Nature 366:365-368(1993).
CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN EXCISION
CC REPAIR OF DNA DAMAGED WITH UV LIGHT, BULKY ADDUCTS, OR CROSS-
CC LINKING AGENTS. ESSENTIAL FOR THE INCISION STEP OF EXCISION-
CC REPAIR.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10275; AAB66928.1; -
DR EMBL; Y07777; CAA69080.1; -
DR EMBL; Z75043; CAA97287.1; -
DR PIR; A23960; A23960.
DR PIR; A29839; A29839.
DR SGD; S0003490; RAD2.
DR InterPro: IPR000513; Exo_N.1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_1.
DR Pfam; PF00867; XPG_1; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS: PR00853; XGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
KW DNA repair; Nuclear protein; Hydrolase; Nuclease; Endonuclease;
KW Magnesium.
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 756 884 I-DOMAIN.
SQ SEQUENCE 1031 AA; 117837 MW; 682D4ACFBDF0F3 CRC64;

Query Match 12.0%; Score 232.5; DB 1; Length 1031;
Best Local Similarity 23.2%; Pred. No. 5.9e-08;
Matches 84; Conservative 67; Mismatches 158; Indels 53; Gaps 10;

OY 54 LTNAGEVTHLQCMFNRTIRLLEAGIKPV---YFDSKPPD-----MKQE---LAKR 101
DB 679 ISKEAEINDSRKVESINKEHR-----KPLIFDISEDEEDNIVEMNIKEQEBDFTEKN 732

```



```

QY 102 YKRDATKDLTEAVEVGDK--DAIEKLSKRTKVTROHNECKRLLRLMGVPEVAPSE 159
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 733 TLTLSAERNVAEENAFVDELFEEQOMKDRDSEVTMDMKIKEYOELLRSRGIPYITAPME 792
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 160 AAEECAALCINDKFAVAVSEMDSLTEGAPRFRLHLDMPSSKRPVMEFVAVLELEL 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 793 AEAQCAELLQMLVDGILITDSDVFLFGGKTKYKNMF--HEKNVYEFYVAESILKLGL 849
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 220 TMDQFIDLCILCGCDYCDISIKIGGOTAKLIRQHSIESILENLNKDRYO----- 270
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 850 DAKNIEELAQLLGSDYTNGLKMGKPVSSIIEVIAEFGLNKNFKDWYNGGFDKRRKQETENK 909
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 271 -----IPEDWYQEARLRFKEPNTLIDPELKTAPDEBELISFLVKDNG 315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 910 FEKDLRKLVLNNEIILDDDPSPVNYDAYMRPEVDHDTTPFWGVPLDMLRSMFTOLG 969
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 316 F-----NEDRVTKALEIKISAKNKSOGRLSEFFKPTATYSAPLKRKETSDDKKAANKK 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 970 WPEHKSDELLPLLRDVAKRRKKKGKQKINEFF--PREYISOD-KKLTSTKRISATGCKLK 1027
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 372 TK 373
      :
DB 1028 KR 1029
      :
RESULT 9
XPG_HUMAN
ID XPG_HUMAN STANDARD: PRT: 1186 AA.
AC P28715.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA-repair protein complementing XP-G cells (Xeroderma pigmentosum
DE group G complementing protein) (DNA excision repair protein ERCC-5).
GN ERCC5 OR XPG OR XPGC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93247645; PubMed=8483504;
RA Scherly D., Nusspikel T., Corlet J., Ucla C., Bairoch A.,
RA Clarkson S.G.;
RT "Complementation of the DNA repair defect in Xeroderma pigmentosum
RT group G cells by a human cDNA related to yeast RAD2.";
RL Nature 363:187-185(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94173288; PubMed=7510366;
RA Shioml T., Hatada Y.-N., Saito T., Shioml N., Okuno Y., Yamazumi M.;
RT "An ERCC5 gene with homology to yeast RAD2 is involved in group G
RT xeroderma pigmentosum.";
RL Mutat. Res. 314:167-175(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94019313; PubMed=8413238;
RA MacInnes M.A., Dickson J.A., Hernandez R.R., Learmonth D., Lin G.Y.,
RA Mudgett J.S., Park M.S., Schaner S., Reynolds R.J., Strniste G.F.,
RA Yu J.Y.;
RT "Human ERCC5 cDNA-cosmid complementation for excision repair and
RT bipartite amino acid domains conserved with RAD proteins of
RT Saccharomyces cerevisiae and Schizosaccharomyces pombe.";
RL Mol. Cell. Biol. 13:6393-6402(1993).
RN [4]
RP SEQUENCE OF 1-88 FROM N.A.
RX MEDLINE=94373034; PubMed=8088606;
RA Samec S., Jones T.A., Corlet J., Scherly D., Sheer D., Wood R.D.,
RA Clarkson S.G.;
RT "The human gene for Xeroderma pigmentosum complementation group G
RT (XPG) maps to 13q33 by fluorescence in situ hybridization.";
RL Genomics 21:283-285(1994).

```

```

RP CHARACTERIZATION.
RX MEDLINE=94266772; PubMed=8206890;
RA O'Donovan A., Scherly D., Clarkson S.G., Wood R.D.;
RT "Isolation of active recombinant XPG protein, a human DNA repair
RT endonuclease.";
RL J. Biol. Chem. 269:15965-15968(1994).
RN [6]
RP CHARACTERIZATION.
RX MEDLINE=94376899; PubMed=8090225;
RA O'Donovan A., Davies A.A., McGys J.G., West S.C., Wood R.D.;
RT "XPG endonuclease makes the 3' incision in human DNA nucleotide
RT excision repair.";
RL Nature 371:432-435(1994).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE=94359802; PubMed=8078765;
RA Habraken Y., Sung P., Prakash L., Prakash S.;
RT "Human Xeroderma pigmentosum group G gene encodes a DNA
RT endonuclease.";
RL Nucleic Acids Res. 22:3312-3316(1994).
RN [8]
RP REVIEW ON VARIANTS XP-G.
RX MEDLINE=99374920; PubMed=10447254;
RA Cleaver J.E., Thompson L.H., Richardson A.S., States J.C.;
RT "A summary of mutations in the UV-sensitive disorders: Xeroderma
RT pigmentosum, Cockayne syndrome, and trichothiodystrophy.";
RL Hum. Mutat. 14:9-22(1999).
RN [9]
RP VARIANT XP-G VAL-792, AND VARIANT ASP-1104.
RX MEDLINE=95038755; PubMed=7951246;
RA Nusspikel T., Clarkson S.G.;
RT "Mutations that disable the DNA repair gene XPG in a Xeroderma
RT pigmentosum group G patient.";
RL Hum. Mol. Genet. 3:963-967(1994).
RN [10]
RP VARIANT XP-G VAL-792.
RX MEDLINE=97250499; PubMed=9096355;
RA Nusspikel T., Lalle P., Leadon S.A., Cooper P.K., Clarkson S.G.;
RT "A common mutational pattern in Cockayne syndrome patients from
RT Xeroderma pigmentosum group G: implications for a second XPG
RT function.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3116-3121(1997).
RN [11]
RP FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
RP EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
RP REPAIR.
RN [12]
RP SUBCELLULAR LOCATION: Nuclear (Probable).
RN [13]
RP DISEASE: COMPLEMENTS THE DEFECT IN XERODERMA PIGMENTOSUM (XP)
RP GROUP G (XP-G). AN AUTOSOMAL RECESSIVE DISEASE CHARACTERIZED BY
RP HYPERSENSITIVITY OF THE SKIN TO SUNLIGHT FOLLOWED BY HIGH
RP INCIDENCE OF SKIN CANCER AND FREQUENT NEUROLOGIC ABNORMALITIES.
RP DISEASE: ALSO INVOLVED IN COCKAYNE'S SYNDROME (CS); AN AUTOSOMAL
RP RECESSIVE DISEASE WHICH IS CHARACTERIZED BY A UV-SENSITIVE SKIN
RP (WITHOUT PIGMENTATION ABNORMALITIES), NEUROLOGICAL DYSFUNCTION
RP DUE TO DEMYELINATION OF NEURONS AND CALCIFICATION OF BASAL GANGLIA
RP (PSYCHOMOTOR RETARDATION, DEAFNESS, OPTIC ATROPHY, RETINAL
RP PIGMENTATION AND HYPERREFLEXES), AND DYSMORPHIC DWARFISM (IMMATURE
RP SEXUAL DEVELOPMENT AND MICROCEPHALY).
RN [14]
RP SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
RP SUBFAMILY.
RN [15]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation
RP at the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).
RN [16]
RP EMBL: X69978; CAA49598.1; -
RP EMBL: D16305; BAA03812.1; -
RP EMBL: I20046; AAC37533.1; -
RP EMBL: X71341; CAA50481.1; -
RP EMBL: X71342; CAA50481.1; JOINED.

```

```

DR PIR: S35993; S35993.
DR MTM; 133530; -.
DR InterPro: IPR000513; Exo_N1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_1.
DR Pfam: PF00867; XPG_1; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
DR DNA repair: DNA-binding; Nuclear protein; Hydroxylase; Nuclease;
KW Endonuclease; Xeroderma pigmentosum; Cockayne's syndrome;
KW Polymorphism; Disease mutation.
KW DOMAIN
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 753 881 I-DOMAIN.
FT DOMAIN 1057 1073 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 254 254 M->V.
FT VARIANT 792 792 A->V (IN XP-G; MILD FORM).
FT VARIANT 1104 1104 /FTID=VAR_007732.
FT VARIANT 1104 1104 /FTID=VAR_007733.
FT VARIANT 1104 1104 H->D.
FT CONFLICT 55 55 /FTID=VAR_007734.
FT CONFLICT 120 122 L->P (IN REF. 2).
FT CONFLICT 126 126 KTA->P (IN REF. 2).
FT CONFLICT 264 266 K->Q (IN REF. 2).
FT CONFLICT 760 760 ROY->Q SSH (IN REF. 2).
FT CONFLICT 796 796 I->F (IN REF. 2).
FT CONFLICT 864 872 I->V (IN REF. 2).
FT CONFLICT 959 959 EG1PTVGV->GNTNGGLC (IN REF. 2).
FT CONFLICT 959 959 R->S (IN REF. 2).
SO SEQUENCE 1186 AA; 133328 MW; 8702EC68E080F1C4 CRC64;

Query Match 11.4%; Score 220.5; DB 1; Length 1186;
Best Local Similarity 21.3%; Pred. No. 4.3e-07;
Matches 94; Conservative 71; Mismatches 149; Indels 127; Gaps 15;

15
15 PRAMEKKESEYFGRIKIVADSMISYQFLIVGRTGMEITNAGEVYTHLQGMFNRTIR 74
635 PRAVEEMELDS-----ESESQSGSFLEV-----QSVISDE-----ELQAEFPEP-- 673
674 -----SKRPSQGEELVGTREGAPAPESSELNDSEKRDVDGDEPOEA--- 717
119 GQKDAIEKL-----SKRTVYTRQHN-----DCKRLL 146
718 -EKDADDSLHEMODINLELEETLESMLAQONSLAKQKQGERIAVYTGQWFLSEQL 776
147 RLKGVVNEVESEAEACALCINDKVFVASEMDSLTFGAPRFLRLHLMDSKKIPVM 206
777 RLFGIRYIOAPMEAEACALCILDLDQTSGITDDSDIMLFGARHYRNFE--NNKFEV 833
207 EEPVAVLELELTMDQFIDLCIGCDSDIKIGGOTAKLIRQ--HGSISILE- 262
834 YIOYVDFHNOGLDKRKLINLAIVLSDTTEGLPYGCYTAHEIINFEFGHG-LEPLKLF 892
263 -----NLNKD--RYQIPEDWPYQEARLLEKPNVTLDIPELKWTA 300
893 SEMMHEAQNKPIRPNRPHDKVKKLRLTQLRPGFPNPAVEAYLTKPVDDSKGSFLMGK 952
301 PDEGLISFLVKNQNGENEDRYTKAI-EKIKSAKNSKSGRLEFEFFPTTISAPLKRRKT 359
953 PDIDKIREFCORYFGWKRRTKTDSELPVVKQDADQOTLRIDSEFPLAQOEKEDAKRIKS 1012
360 S-----DKSKAANK 370
1013 QRLNRVAVTCLMKRKEKKAASE 1033

```

```

RESULT 10
ID XPG_MOUSE STANDARD: PRT: 1170 AA.
AC P35689; 061528; 064248;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA-repair protein complementing XP-G cells homolog (Xeroderma
DE pigmentosum group G complementing protein homolog) (DNA excision
DE repair protein ERCC-5).
GN ERCC5 OR XPG OR ERCC-5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94173288; PubMed=7510366;
RA Shiom T., Harada Y.-N., Saito T., Shiom N., Okuno Y., Yamazumi M.;
RT "An ERCC5 gene with homology to yeast Rad2 is involved in group G
RT xeroderma pigmentosum."
RN Mutat. Res. 314:167-175(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/10; TISSUE=Liver;
RC MEDLINE=96070433; PubMed=7590748;
RA Harada Y.N., Matsuda Y., Shiom N., Shiom T.;
RT "Complementary DNA sequence and chromosomal localization of xpg, the
RT mouse counterpart of human repair gene XPG/ERCC5."
RN Genomics 28:59-65(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=DNA/2;
RC MEDLINE=96359149; PubMed=8703115;
RA Ludwig D.L., Mudgett J.S., Park M.S., Perez-Castro A.V.,
RA MacInnes M.A.;
RT "Molecular cloning and structural analysis of the functional mouse
RT genomic XPG gene."
RN Mamm. Genome 7:644-649(1996).
CC -1- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
CC EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
CC REPAIR (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16306; BAA03813.1; -
DR EMBL; U40796; AAA91039.1; -
DR EMBL; U40795; AAB17885.1; JOINED.
DR EMBL; U39892; AAB17885.1; JOINED.
DR EMBL; U39893; AAB17885.1; JOINED.
DR EMBL; U39894; AAB17885.1; JOINED.
DR EMBL; U39896; AAB17885.1; JOINED.
DR EMBL; U40073; AAB17885.1; JOINED.
DR EMBL; U40431; AAB17885.1; JOINED.
DR EMBL; U40432; AAB17885.1; JOINED.
DR EMBL; U40668; AAB17885.1; JOINED.
DR EMBL; U40669; AAB17885.1; JOINED.
DR EMBL; U40670; AAB17885.1; JOINED.
DR EMBL; U40792; AAB17885.1; JOINED.
DR EMBL; U40793; AAB17885.1; JOINED.
DR EMBL; U40794; AAB17885.1; JOINED.
DR MGD; MGI:103582; ERCC5.
DR InterPro: IPR000513; Exo_N1.
DR InterPro: IPR003584; HNH_2.

```

```

DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HH2.1.
DR SMART: SM00484; XPGI.1.
DR SMART: SM00485; XPGN.1.
DR PROSITE: PS00841; XPG_1.1.
DR PROSITE: PS00842; XPG_2.1.
DR DNA repair; DNA-binding; Nuclear protein; Hydrolase; Nuclease;
KM Endonuclease.
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 752 882 I-DOMAIN.
FT DOMAIN 1049 1065 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT VARIANT 388 388 R -> C (IN STRAIN DBA/2).
FT VARIANT 488 488 S -> R (IN STRAIN DBA/2).
FT VARIANT 688 688 I -> T (IN STRAIN DBA/2).
FT VARIANT 1015 1015 S -> N (IN STRAIN DBA/2).
FT VARIANT 1021 1021 S -> I (IN STRAIN DBA/2).
FT VARIANT 1121 1121 S -> P (IN STRAIN DBA/2).
FT CONFLICT 227 227 N -> M (IN REF. 1).
FT CONFLICT 249 249 N -> M (IN REF. 1).
FT CONFLICT 300 302 VMD -> MDE (IN REF. 1).
FT CONFLICT 313 313 N -> M (IN REF. 1).
FT CONFLICT 320 320 N -> M (IN REF. 1).
FT CONFLICT 399 399 N -> M (IN REF. 1).
FT CONFLICT 408 408 D -> DVQGV (IN REF. 1).
FT CONFLICT 581 581 N -> NSASEVIGPV (IN REF. 1).
FT CONFLICT 795 795 M -> V (IN REF. 1).
FT CONFLICT 1039 1039 A -> AMERFEL (IN REF. 1).
FT CONFLICT 1134 1134 S -> SD (IN REF. 1).
FT CONFLICT 1157 1158 KL -> RR (IN REF. 1).
SQ SEQUENCE 1170 AA; 130864 MW; 4058F07FFD1770ED CRC64;

Query Match 11.3%; Score 220; DB 1; Length 1170;
Best Local Similarity 21.6%; Pred No. 4.6e-07;
Matches 88; Conservative 71; Mismatches 156; Indels 92; Gaps 13;

QY 13 NAKPAKKEKESFYSGRIKIAVDASMSYQFLIVGRTGKMETLTNAGEVYSHL----- 65
DB 633 SAKRPMQPMEMES-----ESESSEDSGSFIEQSVNSSELSQETSSASTHLSKDAEE 684
QY 66 -----OGMNFRTILLEAGIKRVYVFDGKPPDMKKKQELAKRYSRDADATKDLT---EAY 116
DB 685 PRELEGGTSRQTECLIQ-----DSSDIEMEGHRADDDAEDMPMEMODI 730
QY 117 EVGDKDAIEK---LSKRYKVTROHNE-----DCKRLRLMGVPVEAPSEAE 161
DB 731 NLEELDLESNLLAEONSILKAKQKODRIAASVTGOMFLESQELRLRGVPIQAPMEAE 790
QY 162 AECAALCINQKVAFAVASEMDSLTFGAPRFLRHIMDPSSKIFPWE-PDYAKVLELELT 220
DB 791 AQCAMLDSLQSTGTTIDSDIWLFGA---RHVYKNFNNKKFVEYQYDYVSQDLID 846
QY 221 MDGFIDLCILGCGDYCSIKIGTGGATLKLIRQH-GSIESIL-----ENLNK--- 266
DB 847 RNKLINLAYLLGSDYTEGIPVGCYVATILNEFGRCGLDPLKFESEWHKONKKYVAE 906
QY 267 -----DRYQIPEMPYQEARLKEPNVTLIDIPELKTPADEGLISFLVKONG 315
DB 907 NPVDTKKKKLKLQILPQFPNPAVADALRPVDDSRGSLMGKPDVDKIEFQYRFG 966
QY 316 FNEEDVTKATEKI-KSAKNSSGRLSEFFKPLATTSAPLKRKETS 361
DB 967 WNRKKTDESLEYVILKHLNAHOTLRIDSPFR-----LAQGEROD 1005

RESULT 11
EXOT_YEAST STANDARD: PRT: 702 AA.
AC P39875;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)

```

```

DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Exonuclease I (EXO I) (DHSI protein).
EXOT EXOI OR DHSI OR YOR03C OR OR26.23.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RN SEQUENCE OF 220-702 FROM N.A.
RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
RT "Molecular cloning of a gene, DHSI, which complements a drug-
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RA de Haan M., Maarse A.C., Griwell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RA Landt O., Hiesl R., Unseld M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RA Tishkoff D.X., Boerger A.L., Filosi N., Gaida G.M., Bertrand P.,
RA Kane M.F., Kolodner R.D.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RN FUNCTION.
RX MEDLINE=97265410; PubMed=9111347;
RA Florentini P., Huang K.N., Tishkoff D.X., Kolodner R.D.,
RA Symington L.S.;
RT "Exonuclease I of Saccharomyces cerevisiae functions in mitotic
RT recombination in vivo and in vitro.";
RL Mol. Cell. Biol. 17:2764-2773(1997).
CC -1- FUNCTION: 5'->3' double-stranded DNA exonuclease involved in
CC mismatch repair and eventually also in mitotic recombination
CC between direct repeats.
CC -1- COFACTOR: Magnesium or manganese.
CC -1- ENZYME REGULATION: Inactivated by calcium and zinc ions.
CC -1- SUBUNIT: INTERACTS WITH MISMATCH REPAIR PROTEIN MSH2.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S69345; AAC60570.1; -
DR EMBL: X87331; CA60749.1; -
DR EMBL: Z74941; CA99223.1; -
DR EMBL: U86134; AAB47428.1; -
DR PIR: JC2068; JC2068.
DR HSSP: Q29075; TNKL.
DR SGD: S0005559; DHSI.
DR InterPro: IPR000513; EXO_N_I.
DR InterPro: IPR003584; HH2.
DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HH2.1.
DR SMART: SM00484; XPGI.1.
DR SMART: SM00485; XPGN.1.
DR PROSITE: PS00841; XPG_1.1.
DR PROSITE: PS00842; XPG_2.1.1.
KM DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
FT DOMAIN 1 96 N-DOMAIN.

```

FT DOMAIN 114 247 I-DOMAIN.
 FT DOMAIN 501 520 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 537 553 ASP/GRU-RICH (ACIDIC).
 FT DOMAIN 618 625 ASP-RICH (ACIDIC).
 FT SEQUENCE 702 AA; 80162 MW; 7822F6B265DB3AA CRC64;

Query Match 11.2%; Score 218; DB 1; Length 702;
 Best Local Similarity 23.7%; Pred. No. 3.3e-07;
 Matches 99; Conservative 70; Mismatches 176; Indels 72; Gaps 14;

QY 1 MGKIG-LTKLLADNAPKMKQKFESEFGKRIADASMSIYO-----FLIVGRTGMET 53
 1 MGIGLTLQPLKPIQNPVSLR-----YEGVLADGYAMLRHRAACCAVELAMGK----- 50
 DB 54 LTNEGEVTSHLQGMFNFTIRLEAGIKPVYFDGKPPDMKKOELARYSKRDATKDLT 113
 51 -----PDKTLQPIKIRKFSILKTKFVRYPLVPGDADIPVKKSKRKRKE-NKAIA 103
 DB 114 EAV-EVGDKAIEKLSKRTVAVTQHNEDCKRLRLMGVPEVAPSEAEACALCINDK 172
 104 EELMAGCEKKNAMDYFOKCVDTPEMAKCIICYCKLNGIRYIVAFEADSQMYLEQKNI 163
 QY 173 VFAVASEMDSLTFGAPFLRLHMDPSS-----KTIPEVDFVAKYLEEL 219
 164 VGGIISESDLLVFECRLITKLDYEGCELECRDNFTKLEKFPFG-----SL 212
 DB 220 TMDQFIDLCICGDCYCSIKIGGOTALKLRHGSIESILENLNKD-RVQIPEDW--P 276
 213 TNEEITWVCISGCDYTNIGIPKVLITAMKLVRRNTERITLISIORGLMIDPTYINE 272
 QY 277 YOEARRLEKRPV-----TLIDPELKTAPDEGLSLFVKNQGFEDVNTVAIEIK 329
 273 YEAAVLAFQFQFPCIRKRIKIVLSNEIPLYLKDTSKRRLACIGFYIHRETOKQIVH 332
 QY 330 SAKN-----KSSQGRLE--SFEKPTATTSAPLKRKETS-----KTSKAANKTK 373
 DB 333 FDDDDHHLHKLTAQGDLPYDFHOPLANREHKLQLASNSNIEFGKTTVTNSEAKVK 389

RESULT 12
 ID D1N7 YEAST STANDARD; PRT; 430 AA.
 AC 012086;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE DNA-damage inducible protein D1N7.
 GN D1N7 OR D1N3 OR YDR263C OR YD9320B.02C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Murphy L., Harris D., Bartrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: 5'->3' double-stranded DNA exonuclease (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- INDUCTION: By UV light, methyl methane-sulfonate (MMS) or hydroxyurea (HU), and during meiosis.
 CC -1- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X90707; CA62233.1; -
 DR EMBL: 270202; CA94102.1; -
 DR EMBL: 268290; CA92581.1; -
 DR SGD: S0002671; D1N7.
 DR InterPro: IPR000513; Exo_N_1.
 DR InterPro: IPR003584; HHH_2.
 DR InterPro: IPR001532; XPG_1.
 DR Pfam: PF00867; XPG_N_1.
 DR Pfam: PF00752; XPG_L_1.
 DR PRINTS: PR00853; XPGRADSUPER.
 DR SMART: SM00279; HNH2_1.
 DR SMART: SM00484; XPG1_1.
 DR SMART: SM00485; XPGN_1.
 DR PROSITE: PS00841; XPG_1.
 DR PROSITE: PS00842; XPG_2; 1.
 DR DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
 FT DOMAIN 114 247 N-DOMAIN.
 FT SEQUENCE 430 AA; 49034 MW; 2BC23D30832C79E9 CRC64;

Query Match 10.5%; Score 204.5; DB 1; Length 430;
 Best Local Similarity 24.7%; Pred. No. 1.4e-06;
 Matches 68; Conservative 53; Mismatches 129; Indels 25; Gaps 6;

QY 1 MGIGLTLKLLADNAPKMKQKFESEFGKRIADASMSIYO-----FLIVGRTGMETL 54
 1 MGIGLTLQPLKPIQNPVSLR-----YEGVLADGYAMLRHRAACCAVELAMGK----- 51
 DB 55 TNEAGEVTSHLQGMFNFTIRLEAGIKPVYFDGKPPDMKKOELARYSKRDATKDLTE 114
 52 TNEEITWVCISGCDYTNIGIPKVLITAMKLVRRNTERITLISIORGLMIDPTYINE 105
 DB 52 TNK-----YLOFEIKRLQILKRLKIPYVFDGDSLAVANNHETRRKKRLLENIAK 105
 QY 115 AVEVGDKDAIEKLSKRTVAVTQHNEDCKRLRLMGVPEVAPSEAEACALCINDKYE 174
 106 LMSAGNRYNAMEYFOKSVDTPEMAKCIIDYCKLHSIPYVAFEADPQWYLEKMGILQ 165
 DB 175 AVASEMDSLTFGAPFLRLHMDPSSKRIPEVDFVAKYLEEL-----ELTMDQFIDLCIC 231
 DB 166 GIISEDLLVFEGCKTLITKLD--QKALEISKDFSALEPNPGLBLSDQFRNLVCLA 224
 QY 232 GCDYCDISIKIGGOTALKLRHGSIESILENLNK 266
 DB 225 GCDYTGIMKVVYVYAMKIVRYSEMDLILQIER 259

RESULT 13
 ID RA13_SCHPO STANDARD; PRT; 1112 AA.
 AC P28706; O59728;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE DNA repair protein rad13.
 GN RAD13 OR SPB3E7.08C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NCBI_TaxID=4896;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC MEDLINE=93219111; Pubmed=8464724;
 RA Carr A.M., Sheldrick K.S., Murray J.M., Al-Harithy R., Watts F.Z.,
 RA Lehmann A.R.;
 RL "Evolutionary conservation of excision repair in Schizosaccharomycetes
 RT pombe: evidence for a family of sequences related to the

```

RT Saccharomyces cerevisiae RAD2 gene.;
RL Nucleic Acids Res. 21:1345-1349(1993).
RN
RC SEQUENCE FROM N.A.
RC STRAIN-972;
RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Brown D.,
RA Churcher C.M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/CDDB databases.
CC -!- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN EXCISION
CC REPAIR OF DNA DAMAGED WITH UV LIGHT, BULKY ADDUCTS, OR CROSS-
CC LINKING AGENTS. ESSENTIAL FOR THE INCISION STEP OF EXCISION-
CC REPAIR (PROBABLE).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66795; CAA47291.1; -.
DR EMBL: AL023534; CAA19011.1; -.
DR PIR: S22862; S22862.
DR PIR: S30301; S30301.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR001191; Gemini_AL1.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR003903; UIM.
DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00799; Gemini_AL1.1.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
DR DNA repair; Nuclear protein; Hydrolyase; Nuclease; Endonuclease.
FT DOMAIN 1 870 N-DOMAIN.
FT CONFLICT 742 870 D->N (IN REF. 1).
FT CONFLICT 738 743 LKMKR->AQSKRG (IN REF. 1).
FT CONFLICT 738 743 LKMKR->AQSKRG (IN REF. 1).
SQ SEQUENCE 1112 AA; 126328 MW; 7ECF4229D5BF4768 CRC64;

Query Match 10.5%; Score 204.5; DB 1; Length 1112;
Best Local Similarity 21.4%; Pred. No. 4.5e-06;
Matches 88; Conservative 68; Mismatches 148; Indels 107; Gaps 15;

OY 18 MKKQFESEYFGKRIAY--DASMSIYQGLIYGRGMEFLTNEGCVTS----- 63
DB 604 MKKQADDGTVNPLNWSYDRAKMSYV---LASENAKDT-----GDKISSEIDAVLPTLET 655
OY 64 -----HLQGFENRTRLLEAGIKPVYFDGKPPMKKOE---LAKRYSKRDA 108
DB 656 SSPSLSTPDTQKESAPMKGAALSLSSVEPEV--EKLDEEEEMIRAAEEKEKYDRF 713
OY 109 TKDLTE-AVEYGDRAEKL-----SKRTVK-----VTROHNECKRLRLTMGPVVE 155
DB 714 VSELNQRHETERNQENAEKRLKLNKSKSEKRADEAVYQWIKKQCELLRFGILPIYV 773
OY 156 ASEAEACALCTNDKYFAVASDMSLTFGAPRFLR-----HLMDPSKKRP 204
DB 774 AQGEAEACCKSLKELKLVLDGIVTDSVFLFGGTRVYRNMFNONKFEVLYLMDMKREFN 833
OY 205 VAEFPAVAVLELELTMDQFDLCLICGDCYCSIKIGIGOTALKLRQ----- 253
DB 834 VMQMDLIR-----LAHLLGSDYTMGLSRVGVLALELILHERPGDTGLFEERK 879

```

```

OY 254 -----HGSTES---ILENLK--DRYQIPEDMPYQEARLFEKNVTLDIPELKW 298
DB 880 KMFQRLSTGHASKNDVNTVPVKRRKIKLVGKTLIPSEFNPPLVDKAYLHPADDSKOSFOW 939
OY 299 TAPDEGLISFLVKNQGFNEDRVTKAIEKISAKR-----SSQGRLSFPK 345
DB 940 GTPDELEIRGFLMATYGVMSKQRTNEVLPLVTDHMKRQFVGTQSNLQTFE 990

```

RESULT 14

XPG_XENLA STANDARD; PRT; 1196 AA.

```

AC P14629;
DT 01-APR-1993 (Rel. 25; Created)
DT 01-APR-1993 (Rel. 25; Last sequence update)
DT 01-MAR-2002 (Rel. 41; Last annotation update)
DE DNA-repair protein complementing XPG-G cells homolog (Xeroderma
DE pigmentosum group G complementing protein homolog).
GN ERCC3 OR XPG OR XPGC.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93247645; PubMed-8483504;
RA Scherly D., Nusspikel T., Corlet J., Ucla C., Bairoch A.,
RA Clarkson S.G.;
RT "Complementation of the DNA repair defect in Xeroderma pigmentosum
RT group G cells by a human cDNA related to yeast RAD2."
RL Nature 363:182-185(1993).
CC -!- FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE INVOLVED IN DNA
CC EXCISION REPAIR. MAKES THE 3' INCISION IN DNA NUCLEOTIDE EXCISION
CC REPAIR (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. XPG
CC SUBFAMILY.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69977; CAA49597.1; -.
DR PIR: S35994; S35994.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR003584; HNH_2.
DR InterPro: IPR001532; XPG_I.
DR Pfam: PF00867; XPG_I.1.
DR Pfam: PF00752; XPG_N.1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00484; XPG1; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
DR PROSITE: PS00842; XPG_2; 1.
DR DNA repair; DNA-binding; Nuclear protein; Hydrolyase; Nuclease;
FT DOMAIN 1 95 N-DOMAIN.
FT DOMAIN 786 914 I-DOMAIN.
FT CONFLICT 936 952 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT CONFLICT 1079 1095 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 1196 AA; 134206 MW; 1F1CE1891A3C0623 CRC64;

```

```

Query Match 10.3%; Score 200; DB 1; Length 1196;
Best Local Similarity 21.9%; Pred. No. 9.9e-06;
Matches 72; Conservative 64; Mismatches 137; Indels 56; Gaps 11;

OY 88 GKRPDMKKQELAKRYSKRDAIKRLTAEVGVGDKAIEK---LSKRTYKVTROHNE--- 140

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:26:10 ; Search time 26 Seconds
(Without alignments)
2521.734 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 1939
Sequence: 1 MGIGLTKRLADNAPKAKE.....SDKTSKAAANKTKAGKKK 379

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP:REMBL_19:*
2: SP:archaea:*
3: SP:bacteria:*
4: SP:fungi:*
5: SP:human:*
6: SP:invertebrate:*
7: SP:mammal:*
8: SP:mnc:*
9: SP:organelle:*
10: SP:phage:*
11: SP:plant:*
12: SP:rodent:*
13: SP:virus:*
14: SP:vertebrate:*
15: SP:unclassified:*
16: SP:viirus:*
17: SP:bacteriap:*
18: SP:archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1807.5	93.2	380	10	09SX06
2	1313.5	67.7	362	10	065251
3	1069	55.1	382	13	P70040
4	1068	55.1	382	13	P70054
5	1068	55.1	382	13	P70054
6	1063.5	54.8	381	13	Q90YB0
7	1043	53.8	380	11	Q91Z50
8	1040	53.6	380	11	Q91Z50
9	1036.5	53.5	382	11	Q91Z50
10	1009	52.0	378	11	Q91Z50
11	998.5	51.5	385	5	Q91Z50
12	887	45.7	650	5	Q91Z50
13	887	45.7	650	5	Q91Z50
14	648.5	33.4	340	1	Q91Z50
15	636	32.8	343	17	Q91Z50
16	632.5	32.6	343	17	Q91Z50

17	542	28.0	328	17	027670	027670 methanother
18	538.5	27.8	336	17	029975	029975 archaeoglob
19	519	26.8	304	17	0976H6	0976H6 sulfolobus
20	514.5	26.5	401	17	097FY5	097FY5 aeropyrum p
21	506	26.1	302	17	0960U8	0960U8 sulfolobus
22	468	24.1	336	17	094HD4	094HD4 thermoplasm
23	446	23.0	335	17	097B98	097B98 thermoplasm
24	369	19.0	362	17	Q94027	Q94027 halobacteri
25	336	17.3	316	5	096154	096154 plasmodium
26	278	14.3	415	12	091FF5	091FF5 chilo iride
27	267.5	13.8	317	10	0967N8	0967N8 arabidopsis
28	257	13.3	726	5	09VR00	09VR00 drosophila
29	257	13.3	726	5	09VR06	09VR06 drosophila
30	243.5	12.6	1257	5	0917N6	0917N6 drosophila
31	241	12.4	298	13	090CF3	090CF3 red sea bre
32	240	12.4	734	13	09W6K2	09W6K2 xenopus lae
33	235	12.1	497	10	09LEP2	09LEP2 arabidopsis
34	230	11.9	837	11	090Z11	090Z11 mus musculu
35	230	11.9	837	11	0923A5	0923A5 mus musculu
36	227	11.7	800	4	075466	075466 homo sapien
37	227	11.7	803	4	060345	060345 homo sapien
38	227	11.7	803	4	075214	075214 homo sapien
39	227	11.7	846	4	09UNW0	09UNW0 homo sapien
40	227	11.7	846	4	090Q84	090Q84 homo sapien
41	227	11.7	846	4	0961J1	0961J1 homo sapien
42	220.5	11.4	1186	4	09HD59	09HD59 homo sapien
43	219.5	11.3	1479	10	09ATY5	09ATY5 arabidopsis
44	219.5	11.3	1522	10	09LRT2	09LRT2 arabidopsis
45	211	10.9	732	5	024557	024557 drosophila

ALIGNMENTS

RESULT 1	09SX06	PRELIMINARY:	PRT:	380 AA.
ID	09SX06			
AC	09SX06			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	FEN-1.			
GN	OSFEN-1.			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=4530;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RX	MEDLINE=20256470; PubMed=10798612;			
RA	Kimura S., Ueda T., Hatanaka M., Takenouchi M., Hashimoto J.,			
RA	Sakaguchi K.;			
RT	"Plant homologue of flap endonuclease-1: molecular cloning,			
RT	characterization, and evidence of expression in meristematic			
RT	tissues.";			
RL	Plant Mol. Biol. 42:415-427(2000).			
DR	EMBL: AB021666; BAA36171.1; .			
DR	HSSP: 058839; 1A76.			
DR	InterPro: IPR002421; 5_3-exonuclease.			
DR	InterPro: IPR003584; HHH_2.			
DR	InterPro: IPR001532; XPG_1.			
DR	Pfam: PF01367; 5_3-exonuclease; 1.			
DR	Pfam: PF00867; XPG_1; 1.			
DR	Pfam: PF00752; XPG_1; 1.			
DR	PRINTS: PRO0853; XPGRADSUPR.			
DR	SMART: SM00475; 53EXOC; 1.			
DR	SMART: SM00279; Hhh2; 1.			
DR	SMART: SM00484; XPG1; 1.			
DR	SMART: SM00485; XPGN; 1.			
DR	PROSITE: PS00841; XPG_1; 1.			
SO	SEQUENCE 380 AA; 42792 MW; E0148AFA95A7364 CRC64;			

Query Match 93.2%; Score 1807.5; DB 10; Length 380;
 Best Local Similarity 92.1%; Pred. No. 7, 2e-121;
 Matches 350; Conservative 16; Mismatches 13; Indels 1; Gaps 1;

QY 1 MGIGLTKLLADNAPKAMKEOKESFYGRKIAVDASMSIYQFLIVGRTGMETLNEAGE 60
 D 1 MGIGLTKLLADNAPKAMKEOKESFYGRKIAVDASMSIYQFLIVGRTGMETLNEAGE 60
 QY 61 VTSHQGFNFTIRLEAGIKPVYFPGKPPDKKQELAKRYSKRDATKDLTEAVEGD 120
 D 61 VTSHQGFNFTIRLEAGIKPVYFPGKPPDKKQELAKRYSKRDATKDLTEAVEGD 120
 QY 121 KDAIEKSKRTVKTROHNEDECKRLRLMGVPVVEAPEASEACALCINDKFAVASD 180
 D 121 KDAIEKSKRTVKTROHNEDECKRLRLMGVPVVEAPEASEACALCINDKFAVASD 180
 QY 181 MDSLTFGAPRFLRLHMDPSSSKTIPVMEFDYAKVLELELMDQFIDLCTILGCDYCDSTK 240
 D 181 MDSLTFGAPRFLRLHMDPSSSKTIPVMEFDYAKVLELELMDQFIDLCTILGCDYCDSTK 240
 QY 241 GIGGOTALKLIRHGSIESILENLKDRYOIPEDWPYQEARLKEPNVTLDIPELKWTA 300
 D 241 GIGGOTALKLIRHGSIESILENLKDRYOIPEDWPYQEARLKEPNVTLDIPELKWTA 300
 QY 301 PDEGLISFLVKNQFNEDRYTAKIEKISAKNKSOGRLSEFPKRTATSPARKETS 360
 D 301 PDEGLISFLVKNQFNEDRYTAKIEKISAKNKSOGRLSEFPKRTATSPARKETS 360
 QY 361 DRTSKAANKKTK-AGSKKK 379
 D 361 EKPTKAVANKKTKGAGCKK 380

RESULT 2
 ID 065251 PRELIMINARY; PRT; 362 AA.
 AC 065251;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE F2IE10.3 PROTEIN.
 GN F2IE10.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Davidson S., Kohlfling T., David M., O'Brian D.;
 RT "The sequence of A. thaliana F2IE10."
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Washu;
 RT "The A. thaliana Genome Sequencing Project."
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN 13
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Waterston R.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN 14
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Wilson R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF058914; AAC13596.1; -;
 DR HSSP: 058839; 1A76.
 DR InterPro: IPR002421; 5_3_exonuclease.

DR InterPro: IPR000513; Exo_N_I.
 DR InterPro: IPR003584; HHH_2.
 DR InterPro: IPR001532; XPG_I.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF00867; XPG_I; 1.
 DR Pfam: PF00752; XPG_N; 1.
 DR PRINTS: PR00853; XPGADSUPER.
 DR SMART: SM00279; HHH2; 1.
 DR SMART: SM00484; XPG1; 1.
 DR SMART: SM00485; XPGN; 1.
 DR PROSITE: PS00841; XPG_1; 1.
 SO SEQUENCE 362 AA; 41205 MW; A4DD1706C3AD29D CRC64;

Query Match 67.7%; Score 1313.5; DB 10; Length 362;
 Best Local Similarity 72.1%; Pred. No. 1e-85;
 Matches 259; Conservative 30; Mismatches 33; Indels 37; Gaps 3;

QY 1 MGIGLTKLLADNAPKAMKEOKESFYGRKIAVDASMSIYQFLIVGRTGMETLNEAGE 60
 D 1 MGIGLTKLLADNAPKAMKEOKESFYGRKIAVDASMSIYQFLIVGRTGMETLNEAGE 60
 QY 61 VTS-----HLOGMFNFTIRLEAGIKPVYVF 86
 D 61 VTS-----HLOGMFNFTIRLEAGIKPVYVF 86
 QY 87 DGRPPDKKQELAKRYSKRDATKDLTEAVEGDKAIEKLSKRYVKTROHNEDECKRL 146
 D 87 DGRPPDKKQELAKRYSKRDATKDLTEAVEGDKAIEKLSKRYVKTROHNEDECKRL 146
 QY 120 DGRPELLEELQRLMKRYSKRADATLGAIEGKNKEDIEKYSKRYVKTROHNEDECKRL 179
 D 120 DGRPELLEELQRLMKRYSKRADATLGAIEGKNKEDIEKYSKRYVKTROHNEDECKRL 179
 QY 147 RLMGVPVVEAPEASEACALCINDKFAVASDMDSLTFGAPRFLRLHMDPSSSKTIPV 206
 D 147 RLMGVPVVEAPEASEACALCINDKFAVASDMDSLTFGAPRFLRLHMDPSSSKTIPV 206
 QY 180 RLMGVPVVEAPEASEACALCINDKFAVASDMDSLTFGAPRFLRLHMDPSSSKTIPV 239
 D 180 RLMGVPVVEAPEASEACALCINDKFAVASDMDSLTFGAPRFLRLHMDPSSSKTIPV 239
 QY 207 EFDYAKVLELELMDQFIDLCTILGCDYCDSTIKGIGGOTALKLIRHGSIESILENLK 266
 D 207 EFDYAKVLELELMDQFIDLCTILGCDYCDSTIKGIGGOTALKLIRHGSIESILENLK 266
 QY 240 EFEVAKTLEELQRLMDQFIDLCTILGCDYCDSTIRGIGGOTALKLIRHGSIESILENLK 299
 D 240 EFEVAKTLEELQRLMDQFIDLCTILGCDYCDSTIRGIGGOTALKLIRHGSIESILENLK 299
 QY 267 DRYOIPEDWPYQEARLKEPNVTLDIPEL--KMTAPDEGLISFLVKNQFNEDRYTK 323
 D 267 DRYOIPEDWPYQEARLKEPNVTLDIPEL--KMTAPDEGLISFLVKNQFNEDRYTK 323
 QY 300 ERYOIPEDWPYQEARLKEPNVTLDIPEL--KMTAPDEGLISFLVKNQFNEDRYTK 358
 D 300 ERYOIPEDWPYQEARLKEPNVTLDIPEL--KMTAPDEGLISFLVKNQFNEDRYTK 358

RESULT 3
 ID P70040 PRELIMINARY; PRT; 382 AA.
 AC P70040;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
 DE 5' NUCLEASE XFENIA.
 GN XFENI.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 OC NCBI_TaxID=8355;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Bidikova M., Chi E., Wu B., Kim K., Carroll D.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN 12
 RP SEQUENCE FROM N.A.
 RA Bidikova M., Chi E., Wu B., Kim K., Carroll D.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN 13
 RP SEQUENCE FROM N.A.
 RA Bidikova M., Chi E., Wu B., Kim K., Carroll D.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN 14
 RP SEQUENCE FROM N.A.
 RA Bidikova M., Chi E., Wu B., Kim K., Carroll D.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64565; AAB06176.1; -;
 DR EMBL: AF065397; AAD02814.1; -;
 DR HSSP: 058839; 1A76.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR000513; Exo_N_I.
 DR InterPro: IPR003584; HHH_2.
 DR InterPro: IPR001532; XPG_I.

DR Pfam: PF01367; 5_3-exonuclease; 1.
 DR Pfam: PF00867; XPG_I; 1.
 DR Pfam: PF00752; XPG_N; 1.
 DR PRINTS: PR00853; XPGADUPER.
 DR SMART: SM00279; HH2; 1.
 DR SMART: SM00484; XPG1; 1.
 DR SMART: SM00485; XPGN; 1.
 DR Endonuclease.
 KW SEQUENCE 382 AA; 42668 MW; 981DB0EDAD158D57 CRC64;

Query Match 55.1%; Score 1069; DB 13; Length 382;
 Best Local Similarity 55.7%; Pred. No. 2.7e-68;
 Matches 210; Conservative 63; Mismatches 96; Indels 6; Gaps 4;

QY 1 MGIGKGLTKLLADNAPRAKMEQKFESEYFGKRIAVDASMSIYQFLVVGRTGMEITLTNEAGE 60
 DB 1 MGHIGLAKLADVAPAIKEHDIKSYFGKRVAVDASMSIYQFLVAVRODG-NITLONEGE 59
 QY 61 VTSHLGMEFRTIRLMEHGIKRVYVFDGKPPQMKSGELAKRSERRAEKLEAEAGE 119
 DB 60 TTHLMGMEFRTIRLMEHGIKRVYVFDGKPPQMKSGELAKRSERRAEKLEAEAGE 119
 QY 121 KDAIEKLSRTYKVTROHNECKRLRLMGVYVPAEPAEACALCINDKRVAVASED 180
 DB 120 VENIEKFTRLVYKQHNEDCKRLRLMGVYVPAEPAEACALCINDKRVAVASED 179
 QY 181 MDLTFGAPRFLRLHMDSPSSKIPVMEFPAVLELELTMQDFIDLCILGSDYCSIK 240
 DB 180 MDLTFGTPVLLRLHFLASEKKLPIDQEPFLNVIDIGITHEQVLDICILGSDYCSICETIR 239
 QY 241 GIGGQATLKLIRQHSIESILENLKDRYQIPEDMPYQEARLFEKPNVT-LDIPELKWT 299
 DB 240 GIGPKRAIDLIRQHTIDELIDNIDLKRYVPENNLHKAHLFLEPEVDTIDTELKWI 299
 QY 300 APDEELISFLVNDGNEFNDRTKAIERIKSAKNSGGRLESEFFKPTATSAFLKRET 359
 DB 300 EPDEELVAFMCGEKFSEDRIRNGAKKLAKNQGSTQGRIDDFKTVGSVSS-TKRKEA 358
 QY 360 SDRTSRAANKTKAGG 376
 DB 359 E--SKGSAKKRAKTGG 372

RESULT 4

ID 057351 PRELIMINARY; PRT; 382 AA.
 AC 057351;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE FLAP ENDONUCLEASE 1.
 GN FEN1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98204872; PubMed=9535864;
 RA Kim K., Blade S., Matsumoto Y.;
 RT "Involvement of flap endonuclease 1 in base excision DNA repair."
 RL J. Biol. Chem. 273:8842-8848(1998).
 DR EMBL: AF036327; AAB88707.1; -
 DR HSSP: Q58839; 1A76.
 DR InterPro: IPR002421; 5_3-exonuclease.
 DR InterPro: IPR000513; Exo_N_I.
 DR InterPro: IPR003584; HH2.
 DR InterPro: IPR001532; XPG_I.
 DR Pfam: PF01367; 5_3-exonuclease; 1.
 DR Pfam: PF00867; XPG_I; 1.
 DR Pfam: PF00752; XPG_N; 1.
 DR PRINTS: PR00853; XPGADUPER.

DR SMART: SM00279; HH2; 1.
 DR SMART: SM00484; XPG1; 1.
 DR SMART: SM00485; XPGN; 1.
 DR Endonuclease.
 KW SEQUENCE 382 AA; 42668 MW; 3A911F83DC585A59 CRC64;

Query Match 55.1%; Score 1068; DB 13; Length 382;
 Best Local Similarity 55.4%; Pred. No. 3.2e-68;
 Matches 209; Conservative 64; Mismatches 98; Indels 6; Gaps 4;

QY 1 MGIGKGLTKLLADNAPRAKMEQKFESEYFGKRIAVDASMSIYQFLVVGRTGMEITLTNEAGE 60
 DB 1 MGHIGLAKLADVAPAIKEHDIKSYFGKRVAVDASMSIYQFLVAVRODG-NITLONEGE 59
 QY 61 VTSHLGMEFRTIRLMEHGIKRVYVFDGKPPQMKSGELAKRSERRAEKLEAEAGE 119
 DB 60 TTHLMGMEFRTIRLMEHGIKRVYVFDGKPPQMKSGELAKRSERRAEKLEAEAGE 119
 QY 121 KDAIEKLSRTYKVTROHNECKRLRLMGVYVPAEPAEACALCINDKRVAVASED 180
 DB 120 VENIEKFTRLVYKQHNEDCKRLRLMGVYVPAEPAEACALCINDKRVAVASED 179
 QY 181 MDLTFGAPRFLRLHMDSPSSKIPVMEFPAVLELELTMQDFIDLCILGSDYCSIK 240
 DB 180 MDLTFGTPVLLRLHFLASEKKLPIDQEPFLNVIDIGITHEQVLDICILGSDYCSICETIR 239
 QY 241 GIGGQATLKLIRQHSIESILENLKDRYQIPEDMPYQEARLFEKPNVT-LDIPELKWT 299
 DB 240 GIGPKRAIDLIRQHTIDELIDNIDLKRYVPENNLHKAHLFLEPEVDTIDTELKWI 299
 QY 300 APDEELISFLVNDGNEFNDRTKAIERIKSAKNSGGRLESEFFKPTATSAFLKRET 359
 DB 300 EPDEELVAFMCGEKFSEDRIRNGAKKLAKNQGSTQGRIDDFKTVGSVSS-TKRKEA 358
 QY 360 SDRTSRAANKTKAGG 376
 DB 359 E--SKGSAKKRAKTGG 372

RESULT 5

ID P70054 PRELIMINARY; PRT; 382 AA.
 AC P70054;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE XEN1B.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEAD;
 RX MEDLINE=99069415; PubMed=9852084;
 RA Bihlikova M., Wu B., Chi E., Kim K.H., Trautman J.K., Carroll D.;
 RT "Characterization of FEN-1 from Xenopus laevis. cDNA cloning and role
 in DNA metabolism."
 RL J. Biol. Chem. 273:34222-34229(1998).
 DR EMBL: U68141; AAB08478.1; -
 DR HSSP: Q58839; 1A76.
 DR InterPro: IPR002421; 5_3-exonuclease.
 DR InterPro: IPR000513; Exo_N_I.
 DR InterPro: IPR003584; HH2.
 DR InterPro: IPR001532; XPG_I.
 DR Pfam: PF01367; 5_3-exonuclease; 1.
 DR Pfam: PF00867; XPG_I; 1.
 DR Pfam: PF00752; XPG_N; 1.
 DR PRINTS: PR00853; XPGADUPER.
 DR SMART: SM00279; HH2; 1.
 DR SMART: SM00484; XPG1; 1.
 DR SMART: SM00485; XPGN; 1.

SEQUENCE 382 AA: 42865 MW: 10648936A232D460 CRC64:

Query Match 55.1%; Score 1068; DB 13; Length 382;

Best Local Similarity 56.2%; Pred. No. 3.2e-68;

Matches 212; Conservative 60; Mismatches 99; Indels 6; Gaps 4;

```

QY 1 MGIGLTKLADNAPKAMEOKFESYFGKRIAVDASMSIYQFLIYVGRGTGNETLNEAGE 60
  1 MGIGHGLAKLADVAAPAKIKHDIKSYFGKRVAVDASMCYIOFLIYVRODG-NMLONEGE 59
Db
QY 61 VTSHLOGMFKRTIRLLEAGIKPVYVDGKPPDMKKOELAKRYSKDDATKDLTEAVEVD 120
  60 TTSHTLMGMFKRTIRLLEAGIKPVYVDGKPPDMKSGELAKRSEKRAEAEKLEAEFGE 119
Db
QY 121 KDAIEKLSKRTVAVTROHNDCKRLRLMGVVPVPEAPSEAECAALCINXKVFVASED 180
  120 VENIEKNNKLVVTRKOHNECKRLSLMGIPYVDPAPCEAEATCAALYKAGVYAAATED 179
Db
QY 181 MSLTFGAPRFLRLHMDPSSKRIPIVMEFDVAKVLEELTMDQFLDLCILGCDYCSIR 240
  180 MDALTFGTPLVLRHLHASEAKRLPIQEFHLNRFQDIGINHQFVLDLCILGSDYCETIR 239
Db
QY 241 GIGGOTATKLIRHOGSIEILENLKDRYOIPEDMPYOEARLKEPNV-TLDTPELKW 299
  240 GIGPKRAIDLIRHOKTIEILHIDIDILKTPIDPNNMLHKAQQLPEYIDADITELKW 299
Db
QY 300 APDEGLISFLVKGNGFEDRVTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRKE 359
  300 EPDEGLIVAFMCGEKFQSEDRIRNGAKRLAKNKGSTQGRLDPEFKVYGSISST-JKRREV 358
Db
QY 360 SDKTSKAANKTKAGK 376
  359 E--SKGSTKSKSTGG 372
Db

```

RESULT 6

090YB0

ID 090YB0 PRELIMINARY: PRT: 381 AA.

```

AC 090YB0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE FEN-1 NUCLEASE.
GN FEN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsuzaki Y., Adachi N., Koyama H.;
RT "The FEN-1 nuclease is not essential for DNA replication, but is
RT required for maintenance of genomic integrity and base excision repair
RT in vertebrate cells."
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB058602; BAB6507.1; -.
SQ SEQUENCE 381 AA: 43054 MW: 71432ECBCB4E67AA CRC64;

```

Query Match 54.8%; Score 1063.5; DB 13; Length 381;

Best Local Similarity 54.2%; Pred. No. 6.7e-68;

Matches 206; Conservative 67; Mismatches 104; Indels 3; Gaps 3;

```

QY 1 MGIGLTKLADNAPKAMEOKFESYFGKRIAVDASMSIYQFLIYVGRGTGNETLNEAGE 60
  1 MGIGHGLAKLADVAAPAKIKHDIKSYFGKRVAVDASMSIYQFLIYV-RQGAELVONEGE 59
Db
QY 61 VTSHLOGMFKRTIRLLEAGIKPVYVDGKPPDMKKOELAKRYSKDDATKDLTEAVEVD 120
  60 TTSHTLMGMFKRTIRLLEAGIKPVYVDGKPPDMKSGELAKRSEKRAEAEKLEAEFGE 119
Db
QY 121 KDAIEKLSKRTVAVTROHNDCKRLRLMGVVPVPEAPSEAECAALCINXKVFVASED 180
  120 VENIEKNNKLVVTRKOHNECKRLSLMGIPYVDPAPCEAEATCAALYKAGVYAAATED 179

```

```

Db 120 ENIEKFSKRLVKTVOHTDECKRLMLMGIPYVPEAPEAEASCATLVKAGVYAAATED 179
QY 181 MSLTFGAPRFLRLHMDPSSKRIPIVMEFDVAKVLEELTMDQFLDLCILGCDYCSIR 240
  180 MDALTFGTPLVLRHLHASEAKRLPIQEFHLNRFQDIGINHQFVLDLCILGSDYCETIR 239
Db
QY 241 GIGGOTATKLIRHOGSIEILENLKDRYOIPEDMPYOEARLKEPNV-TLDTPELKW 299
  240 GIGPKRAIDLIRHOKTIEILHIDIDILKTPIDPNNMLHKAQQLPEYIDADITELKW 299
Db
QY 300 APDEGLISFLVKGNGFEDRVTKAIEKISAKNKSOGRLSEFPKPTATTSAPLKRKE 358
  300 EPDEGLIVAFMCGEKFQSEDRIRNGAKRLAKNKGSTQGRLDPEFKVYGSISST-JKRREV 359
Db
QY 359 TSDKTSKAANKTKAGK 378
  360 TKGSAKKAKTNSATKAKKK 379
Db

```

RESULT 7

091250

ID 091250 PRELIMINARY: PRT: 380 AA.

```

AC 091250;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIMILAR TO FLAG STRUCTURE-SPECIFIC ENDONUCLEASE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC010203; AAH10203.1; -.
KM Endonuclease
SQ SEQUENCE 380 AA: 42623 MW: 1BE903288B45520D CRC64;

```

Query Match 53.8%; Score 1043; DB 11; Length 380;

Best Local Similarity 53.8%; Pred. No. 1.9e-66;

Matches 205; Conservative 69; Mismatches 99; Indels 8; Gaps 5;

```

QY 1 MGIGLTKLADNAPKAMEOKFESYFGKRIAVDASMSIYQFLIYVGRGTGNETLNEAGE 60
  1 MGIGHGLAKLADVAAPAKIKHDIKSYFGKRVAVDASMSIYQFLIYV-RQGAELVONEGE 59
Db
QY 61 VTSHLOGMFKRTIRLLEAGIKPVYVDGKPPDMKKOELAKRYSKDDATKDLTEAVEVD 120
  60 TTSHTLMGMFKRTIRLLEAGIKPVYVDGKPPDMKSGELAKRSEKRAEAEKLEAEFGE 119
Db
QY 121 KDAIEKLSKRTVAVTROHNDCKRLRLMGVVPVPEAPSEAECAALCINXKVFVASED 180
  120 VENIEKNNKLVVTRKOHNECKRLSLMGIPYVDPAPCEAEATCAALYKAGVYAAATED 179
Db
QY 181 MSLTFGAPRFLRLHMDPSSKRIPIVMEFDVAKVLEELTMDQFLDLCILGCDYCSIR 240
  180 MDALTFGTPLVLRHLHASEAKRLPIQEFHLNRFQDIGINHQFVLDLCILGSDYCETIR 239
Db
QY 241 GIGGOTATKLIRHOGSIEILENLKDRYOIPEDMPYOEARLKEPNV-TLDTPELKW 299
  240 GIGPKRAIDLIRHOKTIEILHIDIDILKTPIDPNNMLHKAQQLPEYIDADITELKW 299
Db
QY 359 TSDKTSKAANKTKAGK 379
  358 PE--PKGPAKKKAKATGGAGK 375
Db

```

RESULT 8

```

Q9JHW7          PRELIMINARY;          PRT;          380 AA.
ID Q9JHW7;
AC Q9JHW7;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE FLAG STRUCTURE-SPECIFIC ENONUCLEASE.
GN FEN-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen D., Gao G., Yang S., Li M., Chen J.;
RT "Cloning and characterization of a rat DNA structure-specific
RT endonuclease (Fen1)."; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF281018; AAF81265.1; -.
DR HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N.1.
DR InterPro; IPR001532; XPG_1.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_1; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; Hnh2; 1.
DR SMART; SM00484; XPG1; 1.
DR SMART; SM00485; XPGN; 1.
DR PROSITE; PS00842; XPG_2; 1.
KW Endonuclease.
SQ
SEQUENCE 380 AA; 42622 MW; 6198CE1F182135 CRC64;

Query Match          53.6%; Score 1040; DB 11; Length 380;
Best Local Similarity 52.6%; Pred. No. 3.1e-66;
Matches 200; Conservative 73; Mismatches 101; Indels 6; Gaps 4;

QY 1 MG1KGLTKLLADNAPKAMKEOKFESYFGKRIADVDAISYOFLLIYVGRGMEITLTNEAGE 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MG10GLAKLADADVAISIRENDIKSYFGKRIADIDASMSIYQFLIIV-RQGDVIVLQNEGE 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 VTSHLQGMFNRTIRLEAGIRPVYVFDGKPPDMKQOELARKYSKDDATKDLTEAVEYGD 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 TTSHLQGMFNRTIRMEENIKIPYIFDGKPPOLKSQAOLAKRSERAEAEKOLQAOEAGA 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 KDAIEKLSKRTVAVRQNHDECKRLRLRMGVVPEVPEASEAEACALCTINDKVFVASED 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 EEEVEKFTKRLVYVKQNHDECKHLISLGMGIYLDAPSEAEASCALAKAGKYAAATFD 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 MDSLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCICGCDYCDISIK 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 MDCLTFSPLVLRHFLAASEKKLPIQEFHLSVLOELGINDQFIDLCILGSDYCESVR 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 GIGGOTALKLIRHOGSISILELNKDRQIPEDMPYQEARLFEKPNVT-LDIPELKWT 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 GIGKRAVDILQHKHGLEELVRLDPSKYRPEENMLHKEARQLFEPEVVPESVELKMS 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 300 APDEGLISFLVKNQGFNEDRYTKAIEKISAKNKSQGRLESPFKPATTSAPLKRET 359
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 EPNEEELVTKMCEKQFSEERIRSGVRLNKSROGSDGRLDFFPKVYGLSS-AKRREP 358
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 360 SDKTSKAANKTKAGGKKK 379
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 E---PKGPARKKAKTGAGCK 375
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
Q9N3T2          PRELIMINARY;          PRT;          382 AA.
ID Q9N3T2;
AC Q9N3T2;
DT 01-OCT-2000 (TREMblrel. 15, Created)

```

```

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HYPOTHETICAL 42.5 KDA PROTEIN.
GN Y47G6A.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Minx P., Graves T., Hawrysko C.;
RT "The sequence of C. elegans cosmid Y47G6A.";
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024791; AAF60653.1; -.
DR HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; Exo_N.1.
DR InterPro; IPR001532; XPG_1.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_1; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; Hnh2; 1.
DR SMART; SM00484; XPG1; 1.
DR SMART; SM00485; XPGN; 1.
KW Hypothetical protein.
SQ
SEQUENCE 382 AA; 42549 MW; EFD60BA0707321EB CRC64;

Query Match          53.5%; Score 1036.5; DB 5; Length 382;
Best Local Similarity 52.3%; Pred. No. 5.6e-66;
Matches 201; Conservative 73; Mismatches 99; Indels 11; Gaps 4;

QY 1 MG1KGLTKLLADNAPKAMKEOKFESYFGKRIADVDAISYOFLLIYVGRGMEITLTNEAGE 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MG1KGLSVIADNAPSAIKVNMKAFFGRTVAIDASMLCTQFLIAYRQDSQ-LQSEDC 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 VTSHLQGMFNRTIRLEAGIRPVYVFDGKPPDMKQOELARKYSKDDATKDLTEAVEYGD 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 TTSHLQGMFNRTIRMEENIKIPYIFDGKPPDMKQOELARKYSKDDATKDLTEAVEYGD 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 KDAIEKLSKRTVAVRQNHDECKRLRLRMGVVPEVPEASEAEACALCTINDKVFVASED 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 VKEAEKFTKRLVYVKQNHDECKHLISLGMGIYLDAPSEAEACALHLYKAGKFGIVTGD 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 181 MDSLTFGAPRFLRHLMDPSSKIPVMEFDVAKVLELELTMDQFIDLCICGCDYCDISIK 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 MDALTFGSVTLRHLFLAVARKIPKEFNLSIALEMKLSVEEFTIDLCILGSDYCGTTR 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 241 GIGGOTALKLIRHOGSISILELNKDRQIPEDMPYQEARLFEKPNVT-LDIPELKWT 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 GVGPKRAVELIQHKNITETLENDQNKYPPPEDMYKRRARELFINEVYKPREVELTKR 299
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 300 APDEGLISFLVKNQGFNEDRYTKAIEKISAKNKSQGRLESPFKPATTSAPLKRET 355
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 EADVEGVYIQLGEMKNFDEERIRNALAKITSKRSYGTRIDSFNGSKVTVCVTAATR 359
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 356 RRTSDKTSKAANKTKAGGKKK 379
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 360 KAEAKAKKGA-----KKGCPK 378

```

RESULT 10
096830 PRELIMINARY; PRT; 378 AA.
AC 096830:
DT 01-MAY-1999 (TREMBlrel..10, Created)
DT 01-MAY-1999 (TREMBlrel..10, Last sequence update)
DE EG:EG0003.3 OR CG8648.
GN Drosophila melanogaster (fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup E.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Godok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jaiswal M., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu G., Zhao Q., Zheng X.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Murphy L., Harris D., Barrell B.;
RT "sequencing the distal X chromosome of Drosophila melanogaster."
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
RA EMBL: AEO03805; AAF57944.1; -.
RA EMBL: AL031863; CAA21330.1; -.
RA HSSP: Q58839; 1A76.
RA FlyBase: FBgn0025832; Feni.
RA InterPro: IPR002421; 5_3_exonuclease.
RA InterPro: IPR000513; Exo_N_I.

```

```

RESULT 11
096830 PRELIMINARY; PRT; 385 AA.
AC 096830:
DT 01-MAY-1999 (TREMBlrel..10, Created)
DT 01-MAY-1999 (TREMBlrel..10, Last sequence update)
DE EG:EG0003.3 OR CG8648.
GN Drosophila melanogaster (fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup E.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Godok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jaiswal M., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu G., Zhao Q., Zheng X.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Murphy L., Harris D., Barrell B.;
RT "sequencing the distal X chromosome of Drosophila melanogaster."
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
RA EMBL: AEO03805; AAF57944.1; -.
RA EMBL: AL031863; CAA21330.1; -.
RA HSSP: Q58839; 1A76.
RA FlyBase: FBgn0025832; Feni.
RA InterPro: IPR002421; 5_3_exonuclease.
RA InterPro: IPR000513; Exo_N_I.

```


[illegible]

Matches	142:	Conservative	62:	Mismatches	107:	Indels	33:	Gaps
QY	19	KEOKFESYPERKATVAVASMSIYOELLIVGRIGMETLTNAEAGVNSHLQGMFNRTIRLLEA	78					
Db	12	KELELEMLYKRAIDALNALITQFSTIRKDDGPIPLMSKRRIITSHLSGLFYRTINLMEA	71					
QY	79	GIRKPVYFEDGKPPDMKQOELARYSKRDDATKDLTEAVEVGDKALIEKLSKRTVYKTRQH	138					
Db	72	GIRKPVYFEDDEPPEFKKLEKRRREAREAEKMEKREALEKEIEEARKYQORATRVNEML	131					
QY	139	NEDCKRLRLMGVAVPEASEAECAALCINDKYFAVASEDMSLTAFGAPREFRLHIMDE	198					
Db	132	IEDAKKLELMGIPITQVASEGEQAAYMAAKGSYYVASASDYSLSLFGAPRLVRNLTIT	191					
QY	199	SSKIP-----VMEFDVAKYLELELTMDQFDILCLCGDY-CDSIKITGG	244					
Db	192	GRRRLPEKNYVEIKELIILE---EYKREKLIRLELALTLVGDVNPGLIKITGL	247					
QY	245	QTALEKLIRHGSIESILENLMD--RYQIPEDWPYQEARLLFKEPNVATIDIPELKATAP	301					
Db	248	KKALEIYVRH-----SKDPLAKFQKOSDVLYALIKKEFFLNIPVT-DNYVLWVRDP	295					
QY	302	DEEGLISFYKNGFNEEDRVYTAIEKIKSAKKKSSQGLSEFFK	345					
Db	296	DEBGIKFLCDEHDFSEERYKNGLERKKAIKSGKSTLSEWFK	339					
RESULT	15							
Q9V0P9								
ID	Q9V0P9	PRELIMINARY:	PRT:	343	AA.			
AC	Q9V0P9:							
DT	01-MAY-2000 (TREMBLrel, 13, Created)							
DT	01-MAY-2000 (TREMBLrel, 13, Last sequence update)							
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)							
DE	DNA REPAIR PROTEIN RAD2 (RAD2).							
GN	PAB1877.							
OS	Pyrococcus abyssi.							
OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.							
OX	NCBI_TaxID=29292;							
RN	11							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=CRSAY;							
RT	"Pyrococcus abyssi genome sequence: Insights into archaeal chromosome structure and evolution."							
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.							
DR	EMBL: AJ248285; CAB49654.1; ..							
DR	HSSP: Q36839; IAT6.							
DR	InterPro: IPR000513; Exo_N_1.							
DR	InterPro: IPR003584; HHH_2.							
DR	InterPro: IPR001532; XPG_1.							
DR	Pfam: PF00867; XPG_I_1.							
DR	Pfam: PF00752; XPG_N_1.							
DR	PRINTS: PRO0853; XPGRADSUPER.							
DR	SMART: SM00279; HHH2_1.							
DR	SMART: SM00484; XPG1_1.							
DR	SMART: SM00485; XPGN_1.							
KW	Complete proteome.							
SO	SEQUENCE 343 AA; 38949 MW; 894A870642B545D CRC64;							
Query Match		32.8%; Score 636; DB 17; Length 343;						
Best Local Similarity		41.6%; Pred. No. 1.6e-37;						
Matches 143; Conservative		58; Mismatches 115; Indels 28; Gaps						
QY	19	KEOKFESYPERKATVAVASMSIYOELLIVGRIGMETLTNAEAGVNSHLQGMFNRTIRLLEA	78					
Db	12	KELELEMLYKRAIDALNALITQFSTIRKDDGPIPLMSKRRIITSHLSGLFYRTINLMEA	71					
QY	79	GIRKPVYFEDGKPPDMKQOELARYSKRDDATKDLTEAVEVGDKALIEKLSKRTVYKTRQH	138					
Db	72	GIRKPVYFEDDEPPEFKKLEKRRREAREAEKMEKREALEKEIEEARKYQORATRVNEML	131					
QY	139	NEDCKRLRLMGVAVPEASEAECAALCINDKYFAVASEDMSLTAFGAPREFRLHIMDE	198					
Db	132	IEDAKKLELMGIPITQVASEGEQAAYMAAKGSYYVASASDYSLSLFGAPRLVRNLTIT	191					
QY	199	SSKIP-----VMEFDVAKYLELELTMDQFDILCLCGDY-CDSIKITGG	244					
Db	192	GRRRLPEKNYVEIKELIILE---EYKREKLIRLELALTLVGDVNPGLIKITGL	247					
QY	245	QTALEKLIRHGSIESILENLMD--RYQIPEDWPYQEARLLFKEPNVATIDIPELKATAP	301					
Db	248	KKALEIYVRH-----SKDPLAKFQKOSDVLYALIKKEFFLNIPVT-DNYVLWVRDP	295					
QY	302	DEEGLISFYKNGFNEEDRVYTAIEKIKSAKKKSSQGLSEFFK	345					
Db	296	DEBGIKFLCDEHDFSEERYKNGLERKKAIKSGKSTLSEWFK	339					
QY	139	NEDCKRLRLMGVAVPEASEAECAALCINDKYFAVASEDMSLTAFGAPREFRLHIMDE	198		</			

```
Db 132 IEDAKKLLQLMGIPYQAPSEGEAOAYMAGGDVYASASODYDSILFGTFRVLVRLNLTIT 191
QY 199 SSKKIPWMEFDV-----AKVLELELTMDOFIDLCILGGDY-CDSIKIGIGOTAL 248
Db 192 GKRMPGKDIYVEIKPELIVLEVELEKELITREKLELAIWGTDYNEGIKIGIPKKAL 251
QY 249 KLIRHGSIESILENLNKD--RYOJPEDMPYQEARRLFKPENVTLDIPELKWTAPDEEG 305
Db 252 EIVK-----YKDBLAKFQROSDVDLYAIKEFFLNPTTDDY-SLKWRPDEEG 299
QY 306 LISPLVNDNGFNEDRYTKALEIKTSANKSSQGRLESEF---KP 346
Db 300 IIRFLCDEHDSERVRNGLERLKAIAKAGOSTLESWPIKKP 343
```

Search completed: November 5, 2002, 16:29:04
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 5, 2002, 09:27.12 ; Search time 1750.69 Seconds
(Without alignments)
16507.481 Million cell updates/sec

Title: US-09-805-311-5

Perfect score: 1381
Sequence: 1 cgaccacgcgtccgcccac.....ttgaaaaaaaaaaaaaaaa 1381

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_bhg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hhg_hum:*
31: em_hhg_inv:*
32: em_hhg_other:*
33: em_hhgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1381	100.0	1381	6	ARI52405	ARI52405 Sequence
2	1350.2	97.8	1463	6	ARI52403	ARI52403 Sequence
3	1348.6	97.7	1541	6	ARI52404	ARI52404 Sequence
4	1334.2	96.6	1478	6	ARI52406	ARI52406 Sequence
5	924	66.9	1354	8	AB021666	AB021666 Oryza sat
6	326.8	23.7	1469	5	XM68141	XM68141 Xenopus lae
7	312.8	22.2	1461	5	XM64563	XM64563 Xenopus lae
8	307	22.2	1357	5	AF065397	AF065397 Xenopus l
9	305.4	22.1	1149	5	AF036327	AF036327 Xenopus l
10	305.4	22.1	2163	10	BC010203	BC010203 Mus muscu
11	304.2	22.1	16086	2	AC026761	AC026761 Mus muscu
12	304.2	22.0	1757	9	HSRAD2	X76711 Homo sapien
13	304.2	22.0	2031	9	BC000323	BC000323 Homo sapi
14	302.8	21.9	138627	2	AP000591	AP000591 Homo sapi
15	302.8	21.9	169053	2	AP000280	AP000280 Homo sapi
16	302.8	21.9	185035	9	AC004770	AC004770 Homo sapi
17	302.8	21.9	196080	9	AC004228	AC004228 Homo sapi
18	302.6	21.9	1144	5	HUMEN1A	L37374 Homo sapien
19	299	21.7	1146	5	AB058602	AB058602 Gallus ga
20	292.4	21.2	39563	9	HSU73629	U73629 Human chrom
21	288.2	20.9	2192	10	AF281018	AF281018 Rattus no
22	279.8	20.3	2033	10	MUSEF1X	L26320 Mouse flap
23	276.2	20.1	6361	10	AF014962	AF014962 Mus muscu
24	274.6	19.9	2114	8	SCYKL13C	Z28113 S.cerevisia
25	274.6	19.9	10660	8	S93804	S93804 BAF1/ABF1/Y
26	260.4	18.9	151865	2	AL356236	AL356236 Homo sapi
27	260.4	18.9	179461	2	AC019177	AC019177 Homo sapi
28	259.4	18.8	2299	8	SPRAD2	X77041 Saccharomyc
29	259.4	18.8	28491	8	SPAC366	Z99167 S.pombe chr
30	258.8	18.7	173256	2	AC073301	AC073301 Homo sapi
31	258.8	18.7	178806	2	AC013557	AC013557 Homo sapi
32	252.6	18.3	37688	2	AL513222	AL513222 Drosophill
33	252.6	18.3	37688	2	DMCEG003	AI031863 Drosophill
34	252.6	18.3	95850	2	AC020003	AC020003 Drosophill
35	252.6	18.3	154965	3	AC004641	AC004641 Drosophill
36	252.6	18.3	160338	3	AC099021	AC099021 Drosophill
37	252.6	18.3	270306	3	AE003805	AC003805 Drosophill
38	241.8	17.5	194322	2	AC024791	AC024791 Caenorhab
39	241.8	17.5	319857	2	AC006782	AC006782 Caenorhab
40	237	17.2	2763	3	AF093702	AF093702 Plasmodiu
41	237	17.2	234112	3	PFMAL4P2	AL035475 Plasmodiu
42	236.6	17.1	1953	3	AF278764	AF278764 Chromosom
43	236.4	17.1	194439	3	CNS078CA	AL590443 Chromosom
44	233.6	16.9	120848	2	AC007801	AC007801 Drosophill
45	159.6	11.6	1023	6	AR062093	AR062093 Sequence

ALIGNMENTS

RESULT 1
LOCUS ARI52405
DEFINITION Sequence 5 from patent US 6232527.
ACCESSION ARI52405
VERSION ARI52405.1 GI:15118455
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1381)
AUTHORS Mahajan,P.B.
TITLE Maize Rad3/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 5 15-MAY-2001;
FEATURES
source Location/Qualifiers
1..1381
BASE COUNT 441 a 269 c 346 g 325 t
ORIGIN

Query Match 100.0%; Score 1381; DB 6; Length 1381;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCCAGCGCTCCGCCACAGCCCGCAGACGAGATGGGCAATCAAGGTTTGAACGAAA 60
 1 CGACCCAGCGCTCCGCCACAGCCCGCAGACGAGATGGGCAATCAAGGTTTGAACGAAA 60
 QY 61 CTGCTGGCGGACAAATGGCGCCCAAGGCGATGAGAGAGAGCAAGTTGAGAGTACTTGGCG 120
 61 CTGCTGGCGGACAAATGGCGCCCAAGGCGATGAGAGAGAGCAAGTTGAGAGTACTTGGCG 120
 QY 121 CGCAAAATGGCGCTGAGCGCCAGCAGTACATGATTTCTGATTTGATGAGAG 180
 121 CGCAAAATGGCGCTGAGCGCCAGCAGTACATGATTTCTGATTTGATGAGAG 180
 QY 181 ACAGGATGGAATCTCTCAAAATGAACTGTGAGTACTGATGATTTGCAAGGAATG 240
 181 ACAGGATGGAATCTCTCAAAATGAACTGTGAGTACTGATGATTTGCAAGGAATG 240
 QY 241 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 300
 241 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 300
 QY 241 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 300
 241 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 300
 QY 301 AAGCTCTCGATATGAGCAACAGAGCTTCTTAAAGATTTCTCAAAAAGAGATGATGCA 360
 301 AAGCTCTCGATATGAGCAACAGAGCTTCTTAAAGATTTCTCAAAAAGAGATGATGCA 360
 QY 361 ACCAAAGATCTGACTGAGGAGTAGAGTAGAGATTAAGATGAGATTTGAGGAGC 420
 361 ACCAAAGATCTGACTGAGGAGTAGAGTAGAGATTAAGATGAGATTTGAGGAGC 420
 QY 421 AAGGAGCTGTAAAGTGCAGAGGAGCAACAGAGATTTGTAAGGCGATTTAAAGCTT 480
 421 AAGGAGCTGTAAAGTGCAGAGGAGCAACAGAGATTTGTAAGGCGATTTAAAGCTT 480
 QY 481 ATGGGGGTTCTGTTTGAAGGCACTTCTGAGAGAGAGAGATGAGAGGCTTTGCG 540
 481 ATGGGGGTTCTGTTTGAAGGCACTTCTGAGAGAGAGAGATGAGAGGCTTTGCG 540
 QY 541 ATAAAGATTAAGGTTTGGCTGCTTCAAGAGATTAAGAGTCTCCCTTACTTTTGGGCT 600
 541 ATAAAGATTAAGGTTTGGCTGCTTCAAGAGATTAAGAGTCTCCCTTACTTTTGGGCT 600
 QY 601 CCACGGTCTCTGCTGATTTAATGAGTCCAAAGTCCAAAGAAATGCTGATGAGAAATT 660
 601 CCACGGTCTCTGCTGATTTAATGAGTCCAAAGTCCAAAGAAATGCTGATGAGAAATT 660
 QY 661 GATTTGCCAAGGTTTGGAGAGGCTTGAACCTCAGACATGACGCTTCAATTTGTTGCG 720
 661 GATTTGCCAAGGTTTGGAGAGGCTTGAACCTCAGACATGACGCTTCAATTTGTTGCG 720
 QY 721 ATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 721 ATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
 QY 781 AAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 781 AAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 QY 841 TATCAAAATTTCTGAGAGCTGCGCTTACCAAGAGCTGAGGCTTGTTCAGAGGCTTAAT 900
 841 TATCAAAATTTCTGAGAGCTGCGCTTACCAAGAGCTGAGGCTTGTTCAGAGGCTTAAT 900
 QY 901 GTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 901 GTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 QY 961 TTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 961 TTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 1021 AAATCTGCCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 1021 AAATCTGCCAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1080

QY 1081 ACCATTCAGCACCCTGTAATAAGGAGAGACTTGGATTAACAGCAGACGCTGCG 1140
 1081 ACCATTCAGCACCCTGTAATAAGGAGAGACTTGGATTAACAGCAGACGCTGCG 1140
 QY 1141 AACAGAAACAAAGGCTGGTGAAGAAAGAAATATCTTGGATGCTTGTATGACACTA 1200
 1141 AACAGAAACAAAGGCTGGTGAAGAAAGAAATATCTTGGATGCTTGTATGACACTA 1200
 QY 1201 CGACTACGAAGAGAGGCGGTGGCGTGAATGCTGCTTGAATTTAATCTCCGTTTGA 1260
 1201 CGACTACGAAGAGAGGCGGTGGCGTGAATGCTGCTTGAATTTAATCTCCGTTTGA 1260
 QY 1261 ACTCAGACTTTGTTGTAAGAGTGGCCGATGTTTGAAGCTGGGCTTACTTGTGTTG 1320
 1261 ACTCAGACTTTGTTGTAAGAGTGGCCGATGTTTGAAGCTGGGCTTACTTGTGTTG 1320
 QY 1321 AAGAGATTGTTGTTACCAAGTACAAACCTTATGCTGTTTGTGAAAAA 1380
 1321 AAGAGATTGTTGTTACCAAGTACAAACCTTATGCTGTTTGTGAAAAA 1380
 QY 1381 A 1381
 1381 A 1381
 Db 1381 A 1381
 1381 A 1381
 RESULT 2
 ARI52403 1463 bp DNA linear PAT 08-AUG-2001
 LOCUS Sequence 1 from patent US 6232527.
 DEFINITION ARI52403
 ACCESSION ARI52403
 VERSION ARI52403.1 GI:15118453
 KEYWORDS
 SOURCE
 ORGANISM
 UNKNOWN.
 REFERENCE
 1 (bases 1 to 1463)
 AUTHORS
 Mahajan, P.B.
 TITLE
 Maize Rad2/FEN-1 orthologues and uses thereof
 JOURNAL
 Patent: US 6232527-A 1 15-MAY-2001;
 FEATURES
 1. 1463
 source
 /organism="unknown"
 BASE COUNT 466 a 292 c 361 g 344 t
 ORIGIN
 Query Match 97.8%; Score 1350.2; DB 6; Length 1463;
 Best Local Similarity 99.48; Pred. No. 0;
 Matches 135; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CGACCCAGCGCTCCGCCACAGCCCGCAGACGAGATGGGCAATCAAGGTTTGAACGAAA 60
 1 CGACCCAGCGCTCCGCCACAGCCCGCAGACGAGATGGGCAATCAAGGTTTGAACGAAA 60
 Db 49 CGCCGCCCGCCACCGCCAGCCGCGCAGACGAGATGGGCAATCAAGGTTTGAACGAAA 108
 49 CGCCGCCCGCCACCGCCAGCCGCGCAGACGAGATGGGCAATCAAGGTTTGAACGAAA 108
 QY 61 CTGCTGGCGGACAAATGGCGCCCAAGGCGATGAGAGAGAGCAAGTTGAGAGTACTTGGCG 120
 61 CTGCTGGCGGACAAATGGCGCCCAAGGCGATGAGAGAGAGCAAGTTGAGAGTACTTGGCG 120
 Db 109 CTGCTGGCGGACAAATGGCGCCCAAGGCGATGAGAGAGAGCAAGTTGAGAGTACTTGGCG 168
 109 CTGCTGGCGGACAAATGGCGCCCAAGGCGATGAGAGAGAGCAAGTTGAGAGTACTTGGCG 168
 QY 121 CGCAAAATGGCGCTGAGCGCCAGCAGTACATGATTTCTGATTTGATGAGAG 180
 121 CGCAAAATGGCGCTGAGCGCCAGCAGTACATGATTTCTGATTTGATGAGAG 180
 Db 169 CGCAAAATGGCGCTGAGCGCCAGCAGTACATGATTTCTGATTTGATGAGAG 228
 169 CGCAAAATGGCGCTGAGCGCCAGCAGTACATGATTTCTGATTTGATGAGAG 228
 QY 181 ACAGGATGGAATCTCTCAAAATGAACTGTGAGTACTGATGATTTGCAAGGAATG 240
 181 ACAGGATGGAATCTCTCAAAATGAACTGTGAGTACTGATGATTTGCAAGGAATG 240
 Db 229 ACAGGATGGAATCTCTCAAAATGAACTGTGAGTACTGATGATTTGCAAGGAATG 288
 229 ACAGGATGGAATCTCTCAAAATGAACTGTGAGTACTGATGATTTGCAAGGAATG 288
 QY 241 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 300
 241 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 300
 Db 289 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 348
 289 TTCAACCGGACATTAAGATTACTGAGAGGGGANTCAAGCGATTTATGTTTGTGAGGC 348
 QY 301 AAGCTCTGATATGAGCAACAGAGCTTCTTAAAGATTTCTCAAAAAGAGATGATGCA 360
 301 AAGCTCTGATATGAGCAACAGAGCTTCTTAAAGATTTCTCAAAAAGAGATGATGCA 360
 Db 349 AAGCTCTGATATGAGCAACAGAGCTTCTTAAAGATTTCTCAAAAAGAGATGATGCA 408
 349 AAGCTCTGATATGAGCAACAGAGCTTCTTAAAGATTTCTCAAAAAGAGATGATGCA 408

QY 361 ACCAAGATCTGACTGAGCGCAGTAGAGATTAAGATCGATTGAAAAATTGAGC 420
|||||
Db 409 ACCAAGATCTGACTGAGCGAGTAGAGATTAAGATCGATTGAAAAATTGAGC 468
QY 421 AAGAGGCTGTAAAGCTCACAAGGCAACACGAAGATTGTAACGGCTATTAGACTT 480
|||||
Db 469 AAGAGGCTGTAAAGCTCACAAGGCAACACGAAGATTGTAACGGCTATTAGACTT 528
QY 481 ATGGGGGTTCTGTGTAGAGGACCTTCTGAAGCAGAAGAGATGTGCAAGCCCTTGGC 540
|||||
Db 529 ATGGGGGTTCTGTGTAGAGGACCTTCTGAAGCAGAAGAGATGTGCAAGCCCTTGGC 588
QY 541 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTGGGCT 600
|||||
Db 589 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTGGGCT 648
QY 601 CCAAGGTTCTCTGTATTAATGATCCAAATTCACAAAGAAATACGTGTGATTTT 660
|||||
Db 649 CCAAGGTTCTCTGTATTAATGATCCAAATTCACAAAGAAATACGTGTGATTTT 708
QY 661 GATGTTCCCAAGTTTTGGAGAGCTTGAACCTCACCATGACCACTCATTTGTCG 720
|||||
Db 709 GATGTTCCCAAGTTTTGGAGAGCTTGAACCTCACCATGACCACTCATTTGTCG 768
QY 721 ATCTGTGTGATGACTATTGTGATAGCATCAAGGATATGCGGGGCAACAGCTCTG 780
|||||
Db 769 ATCTGTGTGATGACTATTGTGATAGCATCAAGGATATGCGGGGCAACAGCTCTG 828
QY 781 AAATTAATTTGTGCAACATGGGCTCATAGAAAGCATTTGGAGATCTTAAATTAAGACGA 840
|||||
Db 829 AAATTAATTTGTGCAACATGGGCTCATAGAAAGCATTTGGAGATCTTAAATTAAGACGA 888
QY 841 TATCAAAATTCCTGAGAGACTGGCTTACCAAGAGCTCGAGCTTGTCAAGAGAGCTTAT 900
|||||
Db 889 TATCAAAATTCCTGAGAGACTGGCTTACCAAGAGCTCGAGCTTGTCAAGAGAGCTTAT 948
QY 901 GTGCAATGATTAATTCCTGAGAGCTAAATGAGTACCTGATGAGAGAGGCTCTAATG 960
|||||
Db 949 GTGCAATGATTAATTCCTGAGAGCTAAATGAGTACCTGATGAGAGAGGCTCTAATG 1008
QY 961 TTCCCTGTAAAGATATATGTTTCAACGAAGATCGGATGACAAAGGCCATAGAGAAATC 1020
|||||
Db 1009 TTCCCTGTAAAGATATATGTTTCAACGAAGATCGGATGACAAAGGCCATAGAGAAATC 1068
QY 1021 AAATTCGCGAAGATTAATTCGTCGAAGGAGACCTCGATCCTTTTCAAGGCAACCTGCC 1080
|||||
Db 1069 AAATTCGCGAAGATTAATTCGTCGAAGGAGACCTCGATCCTTTTCAAGGCAACCTGCC 1128
QY 1081 ACCACATCAGCACCGCTAAAGGGAAGAGACTTCGATTAACAAAGCAAGCAGCTGCG 1140
|||||
Db 1129 ACCACATCAGCACCGCTAAAGGGAAGAGACTTCGATTAACAAAGCAAGCAGCTGCG 1188
QY 1141 AACAGAAAACAAAGGCTGTGGAAGAAAGAAATATCTTGATGCTTGATCAACTA 1200
|||||
Db 1189 AACAGAAAACAAAGGCTGTGGAAGAAAGAAATATCTTGATGCTTGATCAACTA 1248
QY 1201 CGACTACGAAGAGAGCGGTGCGATCCTTCCCTTAACTATTAACTCCCTTTTA 1260
|||||
Db 1249 CGACTACGAAGAGAGCGGTGCGATCCTTCCCTTAACTATTAACTCCCTTTTA 1308
QY 1261 ACCTAGAGCTTTGTTAAAGATTGCGCATGTTCAAGCTGGGGTAAGTATGTTGTTG 1320
|||||
Db 1309 ACCTAGAGCTTTGTTAAAGATTGCGCATGTTTCAAGCTGGGGTAAGTATGTTGTTG 1368
QY 1321 AAGAGATTGTTGACCAAGTAACAAAGCTTATGCTGTTTTTT 1363
|||||
Db 1369 AAGAGATTGTTGACCAAGTAACAAAGCTTATGCTGTTTTTT 1411
RESULT 3
ARI52404
LOCUS ARI52404 1541 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 3 from patent US 6232527.

ACCESSION ARI52404
VERSION ARI52404.1 GI:15118454
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1541)
AUTHORS Mahejan, P. B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 3 15-MAY-2001;
FEATURES
source
1. 1541
BASE COUNT 473 a 308 c 377 g 383 t
ORIGIN
Query Match 97.7%; Score 1348.6; DB 6; Length 1541;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1354; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 CGACCACGCGTCCGCCACAGCCGCCAGACGAGATGGCATCAAGGTTTGACGAAA 60
|||||
Db 43 CGCCGCCGCCACCCGCCACAGCCGCCAGACGAGATGGCATCAAGGTTTGACGAAA 102
QY 61 CTGCTGGCGAGCAATGCGCCCAAGGCGATGAAGGAGCAGAAAGTTGAGAGCTTGGC 120
|||||
Db 103 CTGCTGGCGAGCAATGCGCCCAAGGCGATGAAGGAGCAGAAAGTTGAGAGCTTGGC 162
QY 121 CGCAAAATCGCGTGCAGCGCCAGCATATACAGTTCCTGATTGTATGTTGAAGG 180
|||||
Db 163 CGCAAAATCGCGTGCAGCGCCAGCATATACAGTTCCTGATTGTATGTTGAAGG 222
QY 181 ACAGGATGAAAGCTCTCAAAATGAACTGTAAGTCACTAGTCAATTTGCAAGAAAG 240
|||||
Db 223 ACAGGATGAAAGCTCTCAAAATGAACTGTAAGTCACTAGTCAATTTGCAAGAAAG 282
QY 241 TTCAACCGCAATAAGATTACTGCAAGCGGGAATCAAGCCAGTTATGTTTGTATGC 300
|||||
Db 283 TTCAACCGCAATAAGATTACTGCAAGCGGGAATCAAGCCAGTTATGTTTGTATGC 342
QY 301 AAGCTCTCGATATGAGAAACAAAGAGCTTCTAAAGATCTCAAAAGAGATGATGCA 360
|||||
Db 343 AAGCTCTCGATATGAGAAACAAAGAGCTTCTAAAGATCTCAAAAGAGATGATGCA 402
QY 361 ACCAAGATCTGACTGAGGCACTAGAGTAGAGATTAAGATGCGATTGAAAAATTGAGC 420
|||||
Db 403 ACCAAGATCTGACTGAGGCACTAGAGTAGAGATTAAGATGCGATTGAAAAATTGAGC 462
QY 421 AAGGAGACTGTAAAGGTCACAAAGGCAACAAAGAGATTGTAACGGCTATTAAAGCTT 480
|||||
Db 463 AAGGAGACTGTAAAGGTCACAAAGGCAACAAAGAGATTGTAACGGCTATTAAAGCTT 522
QY 481 ATGGGGGTTCTGTGTAGAGGACCTTCTGAAGCAGAAGAGATGTGCAAGCCCTTGGC 540
|||||
Db 523 ATGGGGGTTCTGTGTAGAGGACCTTCTGAAGCAGAAGAGATGTGCAAGCCCTTGGC 582
QY 541 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTGGGCT 600
|||||
Db 583 ATAAAGCATTAAGTGTTCGCTGTGCTTCAAGATATGACTCCCTTACTTTGGGCT 642
QY 601 CCAAGGTTCTCTGTATTAATGATCCAAATTCACAAAGAAATACGTGTGATTTT 660
|||||
Db 643 CCAAGGTTCTCTGTATTAATGATCCAAATTCACAAAGAAATACGTGTGATTTT 702
QY 661 GATGTTCCCAAGTTTTGGAGAGCTTGAACCTCACCATGACCACTCATTTGATTTGCG 720
|||||
Db 703 GATGTTCCCAAGTTTTGGAGAGCTTGAACCTCACCATGACCACTCATTTGATTTGCG 762
QY 721 ATCTGTGTGATGACTATTGTGATAGCATCAAGGATATGCGGGGCAACAGCTCTG 780
|||||
Db 763 ATCTGTGTGATGACTATTGTGATAGCATCAAGGATATGCGGGGCAACAGCTCTG 822
QY 781 AAATTAATTTGTGCAACATGGGCTCATAGAAAGCATCTTGGAGAAATCTTAATTAAGACGA 840

```

      |||
Db      823  AACTTTTGTGACATGGGTGTCATAGAAAGCATCTTGGAGATCTTAATATAACAGA 882
      |||
Qy      841  TATCAAAATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 900
      |||
Db      883  TATCAAAATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 942
      |||
Qy      901  GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 960
      |||
Db      943  GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 1002
      |||
Qy      961  TTCTGTGTAAGATTAATGTTTCAACGAAGATGGGTGACAAAGCCATAGAGAATC 1020
      |||
Db      1003  TTCTGTGTAAGATTAATGTTTCAACGAAGATGGGTGACAAAGCCATAGAGAATC 1062
      |||
Qy      1021  AATCTGCAAGATTAATGTTGCAAGAGAGCTGAGTCCCTTTTCAAGCACTGCC 1080
      |||
Db      1063  AATCTGCAAGATTAATGTTGCAAGAGAGCTGAGTCCCTTTTCAAGCACTGCC 1122
      |||
Qy      1081  ACCACATGACCGCTAAACAGGAAGAGACTTGGATTAACAAAGCAGCGCTGCG 1140
      |||
Db      1123  ACCACATGACCGCTAAACAGGAAGAGACTTGGATTAACAAAGCAGCGCTGCG 1182
      |||
Qy      1141  AACAGAAACAAAGGCTGTGGAAGAAAGAAATATCTTGGATGTTGATGACACTA 1200
      |||
Db      1183  AACAGAAACAAAGGCTGTGGAAGAAAGAAATATCTTGGATGTTGATGACACTA 1242
      |||
Qy      1201  CGACTACGAAGCAGCGGTGGCTGATCAGTCTGCTTACATTTTAACTCCCTGTTTA 1260
      |||
Db      1243  CGACTACGAAGCAGCGGTGGCTGATCAGTCTGCTTACATTTTAACTCCCTGTTTA 1302
      |||
Qy      1261  ACTGAGAGCTTTGTGTAAGATTTGCCCATGTTTCAAGCTGGGTAGTAGTGTGTTG 1320
      |||
Db      1303  ACTGAGAGCTTTGTGTAAGATTTGCCCATGTTTCAAGCTGGGTAGTAGTGTGTTG 1362
      |||
Qy      1321  AAGAGATTGGTGTACCAAGTAAACAAATCTTATGCTGTTT 1363
      |||
Db      1363  AAGAGATTGGTGTACCAAGTAAACAAATCTTATGCTGTTT 1405
      |||
RESULT 4
ARI52406                               1478 bp      DNA      linear      PAT 08-AUG-2001
LOCUS      ARI52406                               1478 bp      DNA      linear      PAT 08-AUG-2001
DEFINITION  Sequence 7 from patent US 6232527.
ACCESSION  ARI52406
VERSION    ARI52406.1 GI:15118456
KEYWORDS
SOURCE
ORGANISM   Unknown.
            Unclassified.
REFERENCE  1 (bases 1 to 1478)
AUTHORS    Mahajan,P.B.
TITLE      Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL    Patent: US 6232527-A 7 15-MAY-2001;
FEATURES
            location/Qualifiers
            1..1478
            /organism="Unknown"
BASE COUNT 463 a      302 c      365 g      348 t
ORIGIN
Query Match      96.6%; Score 1334.2; DB 6; Length 1478;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1345; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

      |||
Db      181  CGCAAAATGCGCGTGCAGCGCCAGCATGAGCATCTCCAGTTCCTGATATCTAGTTGGAGG 240
      |||
Qy      181  ACAGGATGAAACTCTCAAAATGAGCTGGTGAATGCTACTAGTATTTGCAAGGATG 240
      |||
Db      241  ACAGGATGAAACTCTCAAAATGAGCTGGTGAATGCTACTAGTATTTGCAAGGATG 300
      |||
Qy      241  TTCAACCGGCAATAGATTAATGTTGAGGCGGAATCAAGCCAGTTATGTTTGAATGC 300
      |||
Db      301  TTCAACCGGCAATAGATTAATGTTGAGGCGGAATCAAGCCAGTTATGTTTGAATGC 360
      |||
Qy      301  AAGCTCTGATATGGAAGAACAGAGCTTGTCTAAAGATCTCAAAAAGAGATGCA 360
      |||
Db      361  AAGCTCTGATATGGAAGAACAGAGCTTGTCTAAAGATCTCAAAAAGAGATGCA 420
      |||
Qy      361  ACCAAGATCTGATGAGGAGTAGAGTAGAGATTAAGATGAGTGAAGAAATGAGC 420
      |||
Db      421  ACCAAGATCTGATGAGGAGTAGAGTAGAGATTAAGATGAGTGAAGAAATGAGC 480
      |||
Qy      421  AAGAGAGCTGTAAAGTCAAGAGGCAACAGCAAGATTTGAAAGGCTTTAAGACTT 480
      |||
Db      481  AAGAGAGCTGTAAAGTCAAGAGGCAACAGCAAGATTTGAAAGGCTTTAAGACTT 540
      |||
Qy      481  ATGGGGGTTCTGTGTAGAGGACCTTGTGAAGCAGAGAGAGATGTCAGGCTTTGC 540
      |||
Db      541  ATGGGGGTTCTGTGTAGAGGACCTTGTGAAGCAGAGAGAGATGTCAGGCTTTGC 600
      |||
Qy      541  ATTAAGAGATTAAGTGTTCGCTGTGCTTGAAGATATGAGACTCCCTTACTTTGGGCT 600
      |||
Db      601  ATTAAGAGATTAAGTGTTCGCTGTGCTTGAAGATATGAGACTCCCTTACTTTGGGCT 660
      |||
Qy      601  CCAGGATCTCTGTATTAATGATCCCAAGTCCCAAGAAATACCTGATGGAATTT 660
      |||
Db      661  CCAGGATCTCTGTATTAATGATCCCAAGTCCCAAGAAATACCTGATGGAATTT 720
      |||
Qy      661  GATGTGCCAAGTGTGAGAGCTTTGACATCACCATGACCAAGTATTTGATTTGTC 720
      |||
Db      721  GATGTGCCAAGTGTGAGAGCTTTGACATCACCATGACCAAGTATTTGATTTGTC 780
      |||
Qy      721  ATCCGTGTGATGATGATTTGTGATGATCATTCAAGAGTATGGGGGCAAGAGCTGTG 780
      |||
Db      781  ATCCGTGTGATGATGATTTGTGATGATCATTCAAGAGTATGGGGGCAAGAGCTGTG 840
      |||
Qy      781  AAATTAATTCGTACATGAGGTGTCATAGAGAGATCTTGGAGATCTTAATTAAGACAGA 840
      |||
Db      841  AAATTAATTCGTACATGAGGTGTCATAGAGAGATCTTGGAGATCTTAATTAAGACAGA 900
      |||
Qy      841  TATCAAAATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 900
      |||
Db      901  TATCAAAATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 960
      |||
Qy      901  GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 960
      |||
Db      961  GTCAATTTGATATTCCTGAGAGCTGGCTTACCAAGAAAGCTGAGCGTTGTCAGAGGCTTAAT 1020
      |||
Qy      961  TTCTGTGTAAGATTAATGTTTCAACGAAGATGGGTGACAAAGCCATAGAGAATC 1020
      |||
Db      1021  TTCTGTGTAAGATTAATGTTTCAACGAAGATGGGTGACAAAGCCATAGAGAATC 1080
      |||
Qy      1021  AATCTGCAAGATTAATGTTGCAAGAGAGCTGAGTCCCTTTTCAAGCACTGCC 1080
      |||
Db      1081  AATCTGCAAGATTAATGTTGCAAGAGAGCTGAGTCCCTTTTCAAGCACTGCC 1140
      |||
Qy      1081  ACCACATGACCGCTAAACAGGAAGAGACTTGGATTAACAAAGCAGCGCTGCG 1140
      |||
Db      1141  ACCACATGACCGCTAAACAGGAAGAGACTTGGATTAACAAAGCAGCGCTGCG 1200
      |||
Qy      1141  AACAGAAACAAAGGCTGTGGAAGAAAGAAATATCTTGGATGTTGATGACACTA 1200
      |||
Db      1201  AACAGAAACAAAGGCTGTGGAAGAAAGAAATATCTTGGATGTTGATGACACTA 1260
      |||
Qy      1201  CGACTACGAAGCAGCGGTGGCTGATCAGTCTGCTTACATTTTAACTCCCTGTTTA 1260
      |||
```

[illegible]

Dp	94	TTGGGAGCAATGGCGCCCAAGGCGCATGAGAGGAGAGAAAGTTTCCAGAGCTACTTTCCGGCGG	153
Oy	124	AAATATCGCGTCGAGGCGACGACGTGAGATATACCACTTCCGATGTGTAGTTGGAGGAC	183
Dp	154	CGCATTCGCTCGTGAAGCGGACATGAGATCTTCCAGTCTTCATGTGTGTTGGAGAGG	213
Oy	184	GGCATGGAACCTCTACAAATGAAGCTGTGTAAATGACTACATCATTTTGGACGAATGTG	243
Dp	214	GGGATGGAAACCTCAGCAAGACGCGGTGATGTACCATGATCATTTCAAGGTGTGTC	273
Oy	244	AACCGGACATTAAGATCTGTGAGAGGGGAGCAATCAACCGCTTATGTTTGTGATGAC	303
Dp	274	AACCGGACATTAAGATCTGTGAGAGGGAGCAATCAACCGATATATGTTTGTGATGAC	333
Oy	304	CCCTCGATGTGAGAAAGCAACACACTGCTTAAAGATACCTAAAGATGATGATCAAC	363
Dp	334	CCCTCGATGTGAGAAAGCAACACACTGCTGCAAAAAAGTATCTGAAAGGGAGATGTCCAC	393
Oy	364	AAAGATCTGACGTAGAGGAGTACAGAGTGTAGAGATTAAGATGTGCAAAATTTGGAC	423
Dp	394	AAAGAACTAACAGAGCGAGTACAGAGGATTAAGAGCCATTGTAAATTTTCAGCAG	453
Oy	424	AGGACTGTAAAGTACAAAGGACAGCAACGCAAAATTTGTAAGGGGTATTAAAGCTTAT	483
Dp	454	AGAGCTGTAAAGTACAAAGGACAGCAACGCAAAATTTGTAAGGGGTATTAAAGCTTAT	513
Oy	484	GGGTTCTGTGTGAGAGGACCTTCTGAAAGCAAGACGATATGACAGCTTGTGAT	543
Dp	514	GGGTTCTGTGTGAGAGGACCTTCTGAAAGCAAGACGATATGACAGCTTGTGAT	573
Oy	544	AAAGATGTGTGAGAGGACCTTCTGAAAGCAAGACGATATGACAGCTTGTGAT	603
Dp	574	AAAGATGTGTGAGAGGACCTTCTGAAAGCAAGACGATATGACAGCTTGTGAT	633
Oy	604	CGGTTCTGTGTGATTAAGATGATCAAGGTTCCCAAGAAATATCTGTATGAGATTTGAT	663
Dp	634	CGGTTCTGTGTGATTAAGATGATGATCAAGGTTCCCAAGAAATATCTGTATGAGATTTGAT	693
Oy	664	GGTTCGCAAGGTGTGAGAGGACCTTCTGAAAGCAAGACGATATGACAGCTTGTGAT	723
Dp	694	GGTTCGCAAGGTGTGAGAGGACCTTCTGAAAGCAAGACGATATGACAGCTTGTGAT	753
Oy	724	CTGTGTGATGTGATCTATTGATGATCAATCAAGGATATGAGGGGCAACGCTGTGAAT	783
Dp	754	CTGTGTGATGTGATCTATTGATGATGATCAATCAAGGATATGAGGGGCAACGCTGTGAAT	813
Oy	784	CTTATTCGTCAACATGAGGTGCTCAATGAAAGCACTTGTGAATATCTTAAATGAAGATAT	843
Dp	814	CTTATTCGTCAACATGAGGTGCTCAATGAAAGCACTTGTGAATATCTTAAATGAAGATAT	873
Oy	844	CTTATTCGTCAACATGAGGTGCTCAATGAAAGCACTTGTGAATATCTTAAATGAAGATAT	903
Dp	874	CAATTCCTGAGAGCTGCTTATCAACAAAGCTGACGCTTGTTCAGGAACCCCAATGTT	933
Oy	904	ACATTTGATATTCGTGAGCTTAATAATGAGCTGACCTGATGAGGAGGGTCTCATTAAGTTT	963
Dp	934	ACATTTGATATTCGTGAGCTGATGATGATGATGCTGCCATGAGGAAGGCTGTGTAAGTTT	993
Oy	964	CTGTGTAAGATATGATTTGTTTCAAGCAATCGGTGTGCAAAAGGCTGTGAAAGATCAAT	1023
Dp	994	CTGTGTAAGATATGATTTGTTTCAAGCAATCGGTGTGCAAAAGGCTGTGAAAGATCAAT	1053
Oy	1024	CTGTGTAAGATATGATTTGTTTCAAGCAATCGGTGTGCAAAAGGCTGTGAAAGATCAAT	1083
Dp	1054	TTTCTGCAAGAACTCTTCTCCCAAGSAGGCTGCAATCTCTTTCACAGCTGTGTAG	1113
Oy	1084	ACATTCAGACCCGTAAACGGAAGACCTTGTGGAATTAACGACAAAGCAAGCTGTGAGAC	1143
Dp	1114	ACATTCAGACCCGTCTTAAAGAAAGACCTTGTGGAATTAACGACAAAGCAAGCTGTGAGAC	1173
Oy	1144	AAAGAAACAAAG---GGTGTGCAAAAGCAATATCTGTGATGTGTTGATGACACT	1200

Db	1174	AACAAACAAGGAGCGCCGGCAAGAAGAAATTAATGCATTGTGCATTCATCAACA	1233
OY	1201	CGAC	1204
Db	1234	AGAC	1237
<hr/>			
RESULT 6			
XLU68141	XLU68141	1469 bp	mRNA linear VRI 26-JUL-1998
DEFINITION	Xenopus laevis XENLB mRNA, complete cds.		
ACCESSION	U68141		
VERSION	U68141.1		
KEYWORDS	G1:1549392		
SOURCE			
ORGANISM	African clawed frog.		
Taxonomy	Xenopus laevis		
Phylogeny	Chordata; Craniata; Vertebrata; Euteleostomi;		
Anatomy	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;		
Biological process	Xenopodinae; Xenopus.		
REFERENCE	1 (bases 1 to 1469)		
AUTHORS	Bibikova, M., Wu, B., Chi, E., Kim, K.-H., Grauman, J.R. and Carroll, D.		
JOURNAL	Characterization of FEN-1 from Xenopus laevis. cDNA cloning and		
MEDLINE	1 Biol. DNA metablism		
PUBMED	Chem. 273 (51), 3422-3429 (1998)		
FEATURES			
CDS			
ORIGIN			
1..1469	/organism="Xenopus laevis"		
2 (bases 1 to 1469)	Bibikova, M., Chi, E., Wu, B., Kim, K.-H. and Carroll, D.		
Direct Submission 1996	Biochemistry Univ. of Utah, 50 N. Medical		
Drive, Salt Lake City, UT 84132, USA			
Location/Qualifiers			
1..1469			
/db_xref="taxon:8355"			
/db_xref="taxon:8355"			
111..715-724	/cited_by="Brennan, Brivanlou et al. (1991) Development		
/dev_stage=" tadpole "			
/note="also present in oocytes, as determined by			
sequencing of RT-PCR products"			
function=5'-3' exonuclease and a structure-specific			
endonuclease that removes 5' single-stranded 'flaps';			
implicated in processing of Okazaki fragments during DNA			
replication and in base excision repair; 5' nuclease			
/note="FEN1 homolog; similar to mammalian DnaE IV or			
FEN1 homologs; see also FEN1 and FEN1L, and Og S.			
Genbank Accession Number U64563			
/codon_start=1			
/product="XENLB"			
/protein_id="AAB08478.1"			
/db_xref="gi:1549392"			
AVRDNNLNKLEGGCTTSHMGCPYRITLMERIKPYVDDSKPGSGELARS			
RRARAKNLLEAEAGEVENETKEPKRLKYVRKHNEBECKSLMSG.PYVDPEAS			
ATCAVLAVAGVCYAAATEMDALTEGPRVALDTHLSASEKKLPLOFLNRYVEDIGIS			
NHRDFAPFCILLGSQDYETRIGTGPKRALDIROHKRTIEIINDILKAYIPEMWHLIR			
KEMAFIDDEPFDIVADITELTKTRDEDEGVAFVACGKESFSDINAKKLNNRKQ			
STGDNDLFFRVVSVISDINKNVSEKNSKKNKKTGGTPMGRKNG			
BASE COUNT	461 a	208 c	377 g
ORIGIN	208 c	377 g	345 t
Query Match	23.1%	Score 326.8;	DB 5:
Best Local Similarity	58.1%;	Pred. No. 5,26-74;	Length 1469;
Matches 615;	Conservative 0;	Mismatch 437;	Indels 6;
			Gaps 2;
OY	30	AACGACAGATGGCAGCATGGTTGAAGTCAATCTGCCGCAAGTACGCGCAAGCGAT	89
Db	113	ATGCATAATTCGGAATTCAGCGTTGGCCCAACTTATTCACAGATGCGCGCTGCAAGCTAT	172
OY	90	GAAGACGACAAAGTTTCAGAGCTACTTTCGAGCCCAAAATGCGCGTCGACGACGACATGAG	149

LOCUS	DEFINITION	SEQUENCE	LOCUS	DEFINITION	SEQUENCE	
173	CAAGAGCATGATATATAAAAGTTACTTGGTGGAAAGTGGCCGTGATCTTCTATG	232	173	CAAGAGCATGATATATAAAAGTTACTTGGTGGAAAGTGGCCGTGATCTTCTATG	232	
Oy	150	CATATATACATCTTCGATTTCTAGTGGAGCACAGCCGTGAACACTCTCAAAAGTAC	209	Oy	150	CATATATACATCTTCGATTTCTAGTGGAGCACAGCCGTGAACACTCTCAAAAGTAC
Db	233	CATTTCACATCTCTCTATCGCTGCACAGATAGGCA--ATATGCTGCACATGAGCA	289	Db	233	CATTTCACATCTCTCTATCGCTGCACAGATAGGCA--ATATGCTGCACATGAGCA
Oy	210	TGTGTAACTACTACTACTATTTTTCAGAGGAATGTTCACCGACATTAATATGTGAGC	269	Oy	210	TGTGTAACTACTACTACTATTTTTCAGAGGAATGTTCACCGACATTAATATGTGAGC
Db	290	AGGCAACACACACACATTAATGATAGGATATGTTTTCACCTGATACCTTGATGAGCA	349	Db	290	AGGCAACACACACACATTAATGATAGGATATGTTTTCACCTGATACCTTGATGAGCA
Oy	270	GGGATCAACACCCAGTTATAGTTTGTGGAGAACCCCTCGATATGAGAAACAGAGCT	329	Oy	270	GGGATCAACACCCAGTTATAGTTTGTGGAGAACCCCTCGATATGAGAAACAGAGCT
Db	350	TGCATCTCAAAACCCGTTTATGTGTGTTCAGTGGAAACCCGCTCAATGAAAGTCAGTAC	409	Db	350	TGCATCTCAAAACCCGTTTATGTGTGTTCAGTGGAAACCCGCTCAATGAAAGTCAGTAC
Oy	330	TGCTTAAAGTACTCAAAAGATCATGTCAGACCAAAATCTCATGAGGCAAGTAAAGCT	389	Oy	330	TGCTTAAAGTACTCAAAAGATCATGTCAGACCAAAATCTCATGAGGCAAGTAAAGCT
Db	410	AGCCAAAGGAGTAGAGCTTAGGCGAGAGGCAAGAAAGTTGCTTCAAGCTCCGAGAGAGC	469	Db	410	AGCCAAAGGAGTAGAGCTTAGGCGAGAGGCAAGAAAGTTGCTTCAAGCTCCGAGAGAGC
Oy	390	AGAGAGTAAGATGTGGTATTTGAAATTTGACACAAAGAGCTGTAAAGTGTCAAGGCAAC	449	Oy	390	AGAGAGTAAGATGTGGTATTTGAAATTTGACACAAAGAGCTGTAAAGTGTCAAGGCAAC
Db	470	TGACAGATGTGGAAATGATACAAATTCATTAATATATGAGGAGCTGTGTAAGCAACAGC	529	Db	470	TGACAGATGTGGAAATGATACAAATTCATTAATATATGAGGAGCTGTGTAAGCAACAGC
Oy	450	CACAGCAATTTGTAAAGCGCTTAAAGATTTATGAGGAGCTGTGTAAGCAACAGCCTTC	509	Oy	450	CACAGCAATTTGTAAAGCGCTTAAAGATTTATGAGGAGCTGTGTAAGCAACAGCCTTC
Db	530	CATATCAATGTGAGAAATTTCTTACCTATGGTATTTCTTTATGTGATGACACCCCTG	589	Db	530	CATATCAATGTGAGAAATTTCTTACCTATGGTATTTCTTTATGTGATGACACCCCTG
Oy	510	TGAAGCAGACAGAAATGTGACGCCCTTGGATTAACGATTAAGATGTTCCTGTGCTTTC	569	Oy	510	TGAAGCAGACAGAAATGTGACGCCCTTGGATTAACGATTAAGATGTTCCTGTGCTTTC
Db	590	TGAGATGTGACACATGTGCTGCTTACTATTAAGAGAAAGATGTATCTCTACAGAC	649	Db	590	TGAGATGTGACACATGTGCTGCTTACTATTAAGAGAAAGATGTATCTCTACAGAC
Oy	570	AGAAAGTTGTGACTCCTTACTTTTGGGGGCTGACCGCTTCTCTCTATTAATGATATGC	629	Oy	570	AGAAAGTTGTGACTCCTTACTTTTGGGGGCTGACCGCTTCTCTCTATTAATGATATGC
Db	650	TGAAGACATGTGATGCTCTGACCTTTGGCATTTGCGATGCTGTATGGCCATCAGCAAG	709	Db	650	TGAAGACATGTGATGCTCTGACCTTTGGCATTTGCGATGCTGTATGGCCATCAGCAAG
Oy	630	AAGCTTCACAAATATACCTGTGATTAAGATTTGATGATGCTCAAGTGTTCAGGACACTGA	689	Oy	630	AAGCTTCACAAATATACCTGTGATTAAGATTTGATGATGCTCAAGTGTTCAGGACACTGA
Db	710	TGAAGCTAAAGCTGCTTATTCAGAGGATTCATCTAAACCGGATTTTTCAGAAATCTGG	769	Db	710	TGAAGCTAAAGCTGCTTATTCAGAGGATTCATCTAAACCGGATTTTTCAGAAATCTGG
Oy	690	ACTCACATGTGACAGGTTATGATTTGTGATCTGTGTGTGATGTGACATTTGTGATG	749	Oy	690	ACTCACATGTGACAGGTTATGATTTGTGATCTGTGTGTGATGTGACATTTGTGATG
Db	770	CATCATATCACACAAATTTGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	829	Db	770	CATCATATCACACAAATTTGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
Oy	750	CATCAAAAGTATGTGGGGGCAACAGCTGTGAAACTTTGTGTCAATGTGGGTCCATAGA	809	Oy	750	CATCAAAAGTATGTGGGGGCAACAGCTGTGAAACTTTGTGTCAATGTGGGTCCATAGA
Db	830	AATTCGATGATGTGACCCCAAAAGCATTGACATGATATGCGCAGCATTAAGCATTTGA	889	Db	830	AATTCGATGATGTGACCCCAAAAGCATTGACATGATATGCGCAGCATTAAGCATTTGA
Oy	810	AAGCTTCAGCAATTTTAAAGAAAGATGATTAATTCCTGATGATGAGGCTGCTTACA	869	Oy	810	AAGCTTCAGCAATTTTAAAGAAAGATGATTAATTCCTGATGATGAGGCTGCTTACA
Db	890	AGAGATCATGTGATTAACATTGATGATCTCAAGAAATATCCCATTCAGAGAAATCGCTCCAA	949	Db	890	AGAGATCATGTGATTAACATTGATGATCTCAAGAAATATCCCATTCAGAGAAATCGCTCCAA
Oy	870	AGAGATGTGACGCTGTGTCAAGGCTTAATGTGACA--TTGAGATTTCTGTACACTAA	926	Oy	870	AGAGATGTGACGCTGTGTCAAGGCTTAATGTGACA--TTGAGATTTCTGTACACTAA
Db	950	GGAGGCGAGCGAGCTTTCTGTGAAACAGAGATGATTAATGATCTGATATCTGATACCTA	1009	Db	950	GGAGGCGAGCGAGCTTTCTGTGAAACAGAGATGATTAATGATCTGATATCTGATACCTA
Oy	927	ATGAGCTACACCTGATGAGAGGCTGTCAATGTTCTGTGTAAGAAAGTAAATGTTGTTCA	986	Oy	927	ATGAGCTACACCTGATGAGAGGCTGTCAATGTTCTGTGTAAGAAAGTAAATGTTGTTCA
Db	1010	GTTGAGCTACCTGATGAGAGGCTGTGATGAGGCTGCTTCTATGTGTGAGGAGAACTGATC	1069	Db	1010	GTTGAGCTACCTGATGAGAGGCTGTGATGAGGCTGCTTCTATGTGTGAGGAGAACTGATC
Oy	987	CGAAGTACGCTGATCAAAAGCGATGAGACATTAATGATTAATGATTAATGATGCTCCA	1046	Oy	987	CGAAGTACGCTGATCAAAAGCGATGAGACATTAATGATTAATGATTAATGATGCTCCA
Db	1070	TGAAGATCATGTGATGATGAGGCGCAAAATTTGTGATGATTAATGATTAATGATGATG	1129	Db	1070	TGAAGATCATGTGATGATGAGGCGCAAAATTTGTGATGATTAATGATTAATGATGATG
Oy	1047	AGAGAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1084	Oy	1047	AGAGAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Db	1130	GGGCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1167	Db	1130	GGGCGGCTGTGATGATGATG

VERSION U64563.1 GI:1490869
 KEYWORDS African clawed frog.
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 1461)
 Bibikova, M., Wu, B., Chi, E., Kim, K. H., Trautman, J. K. and Carroll, D.
 TITLE Characterization of FEN-1 from Xenopus laevis. cDNA cloning and
 JOURNAL role in DNA metabolism
 J. Biol. Chem. 273 (51), 34222-34229 (1998)
 MEDLINE 99069415
 PUBMED 9852084
 REFERENCE 2 (bases 1 to 1461)
 Bibikova, M., Chi, E., Wu, B., Kim, K. H. and Carroll, D.
 TITLE Direct Submission
 JOURNAL Submitted (19-JUN-1996) Biochemistry U. of Utah Med. Sch., 50 N.
 Medical Dr., Salt Lake City, UT 84132, USA
 FEATURES
 source
 1. 1461
 /organism="Xenopus laevis"
 /db_xref="taxon:8335"
 11. 1259
 /function="implicated in processing Okazaki fragments
 during DNA replication and in base excision repair;
 specifically removes single-stranded 5' sequence"
 /note="similar to FEN-1, DnaseIV, Mfl, RAD27, RTH1, and
 rad2"
 /codon_start=1
 /product="5' nuclease xFEN1a"
 /protein_id="AAB06176.1"
 /db_xref="GI:1490870"
 /translation="MGHIGLKLINDVAPALIKEDIRKSYGRKVAADASMCYIPLI
 AVRODNTLNEGEFTSHMGPEYRTIMYEGKTPYVDEGPPROMKSGELAKRSE
 RREBAKLLEAEBAEVEENITKTKRLVKYTKHNECKKLHLIEMLPYDAFCEAE
 ATCAALVAKAGYAAATEDMDALFTGTPLVLRHLHLSAKKLPIQEFHLNVLDIGI
 THEQFDLILGSDYCEITRIGIKPKRAIDLIRKKTIDELIDIRKYPVPMNMIH
 KEAKHLFEPVVDIDITELKWIPEDEGLVAFMGCEQFEDIRNGAKYKPKNNROG
 STQGRLDPDFFKYSSVSKRFEAKSKAKKATGCTPGAKPKRK"
 BASE COUNT 454 a 293 c 365 g 349 t
 ORIGIN
 Query Match 22.7%; Score 312.8; DB 5; Length 1461;
 Best Local Similarity 53.0%; Pred. No. 2.3e-70;
 Matches 715; Conservative 0; Mismatches 627; Indels 6; Gaps 2;

Db 468 GTGGAGAACATTGAGAGTCTCAATAAAGACTTGTCAAGTTACCAAGCAACCAATGAG 527
 Qy 457 GATTGTAAGGCGCTATTAGACTTATGGGGTCCCTGTTGAGAGCAGCTTCGAAGCA 516
 Db 528 GAGTGGCAAAATAATTACTTACCTTAACCTTAATGGGCAATCTCTATGATGGATGCAACCTGGAAGCT 587
 Qy 517 GAAGCAGAAATGTGACGCCCTTGTGCATAAAGCATTAAGAGTGTGCGTGTGCTGTCATGAAGAT 576
 Db 588 GAGCGCACTGTGTGCTCTTATTAACACAGGAAAGTGTGCTGACGACAGCAAGAAAG 647
 Qy 577 ATGAGCTCCCTTACTTGTGGGGCTCCACGGTTCTCTGCTATTTAATGATCCAAATTC 636
 Db 648 ATGATGACCCCTGACCTTGGCACCCCTGCTCTTACGACCTCAACAGCAAGCAAGCT 707
 Qy 637 AAGAAATATCTGTGATGAATTTGATGTGCAAGGTTTGGAGAGCTTGAAGCTCAC 696
 Db 708 AAAAAGCTCCATATTCAGSAGTTCATTAACCGTGTATTCAGGACATAGATATCACA 767
 Qy 697 ATGGACCACTTCAATGATTTGTGATCTCTGTGTGATGTGACTATTGTGATGATCAAA 756
 Db 768 CATTGAACAATTTGTGATCTCTGTATCTTGTCTCGAAGTGTACTGTGAACCATCCGT 827
 Qy 757 GGTATGGGGGCAAAAGCTCTGAACCTTATGCTCAACATGGTCCATAGAAAGCATC 816
 Db 828 GGCATTGAGCAACCAAAAGAGCCATTGACTGATCGTACGATTAAGACATTGAGAGATC 887
 Qy 817 TTGGAGATCTTAATTAAGACAGATATCAATTCCTTGAGAGCTGCGCTTACCAAGAGCT 876
 Db 888 ATTGATTAACATTTGACCTCAAGAAATACCCCTCGGAAATCGCGCAACAGAGAGCA 947
 Qy 877 CGAGCGCTTGTTCAGAGAGCGCTAA--TGTCACTATGATATTCCTGACCTTAATATGACT 933
 Db 948 AATCATCTTTCTTGTGAACCAAGATAGTGCAGACTGACATCACGAGAGCTGAATGATC 1007
 Qy 934 GCACCTGATGAGAGAGGCTCTCAAGTTCCGCTGTAAGAAAGTAATGGTTTCAACGAAGAT 993
 Db 1008 GACCCGATGAAGAAGATGTTGTGCTTCATGTGTGGGGAAGACAGTTAGTAAGAT 1067
 Qy 994 CGGATGACAAAGGCCATAGAGAAGATCAATCTGCCAAGTAATATGCTGCAAGAGAGA 1053
 Db 1068 CGCATACGCAACGGGGCCAAAGTTGGCAAGAACCGGCAAGGAGACACAGGCGCG 1127
 Qy 1054 CTCGAGTCTTTTTCAGAGCACTGCCACACATAGACAGCCGCTTAAGAGAGAACT 1113
 Db 1128 CTGATGACTTTTTCAGAGTACCGGATCCGTTAGCTTAATTAAGAAAGAGGAGAA 1187
 Qy 1114 TCGATTAACAAGCAAGGACAGCTCGAACAAGAAACAAAGGCTGTGTGAAGAAAGAAA 1173
 Db 1188 TCCAAAGATCCGCTTAAGAAAGAAAGCAAAACTGTGAGTACCCTGCGAAGATTCAG 1247
 Qy 1174 TATCTGTGATGCTGTATGATACACTACGACACGCAAGACAGCGGTGCGATGACTTC 1233
 Db 1248 AGGGGTAATGAATGGGCAATTTTAATCTGATGCAATGAATTTTATMACCGCTTAAGGCTTT 1307
 Qy 1234 GCTTATTAATTAATCTCCCTGTTTAACCTAGAGCTTTGGTAAAGTTGCGCCATGTT 1293
 Db 1308 GACAAATATGCGAGTCTTAAATAAATAAATAATTTGTAATAATACCTTATCTTAAG 1367
 Qy 1294 CAACTGGGGTAAGTATGTTGTGTTGAAGATATGGTGCACAGTAATACAAACTATC 1353
 Db 1368 TGGGTAACTTTAATGCTAATGTTGTTTGTATTAATAATTAAGATTAACCAAGTTC 1427
 Qy 1354 GCTGTTTTTGAAGAAAAAAGAAAAA 1381
 Db 1428 TGTCTGTGTCATGAAGAAAAAAGAAAAA 1455
 RESULT 8
 AF065397 1357 bp mRNA linear VRT 05-JAN-1999
 LOCUS
 DEFINITION Xenopus laevis flap endonuclease 1 (FEN1) mRNA, complete cds.
 ACCESSION AF065397
 VERSION AF065397.1 GI:4106357


```

/clone_lib="NCI_GAP_Mam6"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
206..1348
CDS
/codon_start=1
/product="Similar to flag structure-specific endonuclease"
/protein_id="AAH10203.1"
/db_xref="GI:16307328"
/translatability="HIGH"
/translation="MGIHGLKLIADVAPSAIRENDIKSYFGKVAIDASMSIYOLLI
ATROGDIVLONBEGTSTHLMGFTRTIRIMENIKVIVYFDGKPOLKSGELAKSE
RAAEAKOLOQAQEAEMEEVEKFTKRLVYTKQHNDECKHLISLPIYLDPSAE
ASCAALAKQAKVYAAATMDMLTFFGSPVLMRHLTJASBAKLLPQEHLSRVQLEL
NOEVPDLCILISDYCESIRIGIKRAVDLIQKHSIEELVRLDPSKRPVENMILH
KEAOLFLPEVLQDESEYELKSENEELVYKFMCGKROKSEERIRGKVRRLSKSROG
STOGRLDPEFVYTGSLSAKKRKEPEPKAKKKKTKGAGKFRGK"
BASE COUNT      540 a      528 c      619 g      476 t
ORIGIN
Query Match      22.1%; Score 305.4; DB 10; Length 2163;
Best Local Similarity 56.3%; Pred. No. 2e-68;
Matches 642; Conservative 0; Mismatches 481; Indels 18; Gaps 3;
QY 37 ATGGGCAATCAAGGCTTGAACGAACTGCTGGGGGAGACAAATGCGCCCAAGCCATGAGAG 96
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 206 ATGGGAATTCAGCGCCTTGCCAACTAATGCTGATGTGCCCCCAAGTCCCTGAG 265
QY 97 CAGAAATTGAGAGCTACTTCGGCCCGCAAAATGCGCGTGCAGCCGACATGAGATATAC 156
    || || || || || || || || || || || || || || || || || || || || || ||
Db 266 AATGACATCAAGAGCTACTTGTGTCGCAAAAGTGCATGATGCTCCATGAGCATCTAC 325
QY 157 CAGTTCCTGATGTTGTAAGAGACGAGCATGGAACCTCACAATAAGAGCTGAGTGA 216
    ||||| ||||| ||||| || || || || || || || || || || || || || || || ||
Db 326 CAGTTCCTGATGTTGTTGCTGCTCAG--GTTGGGAGTGTCTCAGACAGAGAGGAGT 382
QY 217 GTCACATGATTTGCAAGAGATGTTCAACCGGACATTAAGTACTGAAACCGGAATC 276
    || || || || || || || || || || || || || || || || || || || || || ||
Db 383 ACCACCGACCACTGATGATGTCATCTACCGTACCGATGATGATGAGATGAGTGCATC 442
QY 277 AAGCCATGATGTTGTTGTAAGAGAGCCGCTGATGATGATGATGATGATGATGATG 336
    ||||| ||||| ||||| ||||| || || || || || || || || || || || || || ||
Db 443 AAGCCTGTGATGCTCTTTGATGTCACCAACACACAGCTAAGTCAAGGAGTGGCCAG 502
QY 337 AGATATCTCAAAAAGAGATGATGACCAAAAGATCTGATGAGGAGTGAAGTGAAGAT 396
    || || || || || || || || || || || || || || || || || || || || || ||
Db 503 CGCAGTGAAGAGCGCGCGCGAGCTGAGAAAGCAACTCAGACGAGGCTGAGGATG 562
QY 397 AAGATGCGATGAAAAATGACAGAGAGAGCTGTAAGTGTCAACAAGGCAACACAGCA 456
    || || || || || || || || || || || || || || || || || || || || || ||
Db 563 GAGGAGAGGTTGAGAGATTCACCAAGAGAGCTGTGAAGGTCAACAGACACATGAT 622
QY 457 GATTGTAACGGCTATTAAGACTTATGAGGAGTCTGTTGTAGAGGACACTTGTGAACA 516
    || || || || || || || || || || || || || || || || || || || || || ||
Db 623 GAGTGAACAAACCTGCTGATGAGGAGTCTGATGATGATGATGATGATGATGATG 682
QY 517 GAGCCAAAGTGTGAGCGCTTGTGATTAAGAGTGTGAGTGTGAGTGTGAGTGTGAG 576
    || || || || || || || || || || || || || || || || || || || || || ||
Db 683 GAGGCAAGCTGTGTGCTGCGCAAGGCTGCAAAAGTGTATGTGCTGCGGCAAGGAG 742
QY 577 ATGAGACTCCCTTACTTTGGGGCTCCACAGTTCCCTTGTGATTTAATGATCAAGTTCC 636
    ||||| ||||| ||||| || || || || || || || || || || || || || || || ||
Db 743 ATGAGACTCCCTTACTTTGGGCTCCCGGCTGATGATGATGATGATGATGATGATG 802
QY 637 AAGAAATFACCTGTGATGAGATTTGATGTTGCCAAGTTTGGAGAGAGTTTAATCACC 696
    ||||| || || || || || || || || || || || || || || || || || || || ||
Db 803 AAGAGCTGCGCATCCACAGATTCATGAGCGCGCTGCTGAGAGAGTGGGCTGTGAC 862
QY 697 ATGAGACCAATTCATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 756
    || || || || || || || || || || || || || || || || || || || || || ||
Db 863 CAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 922
QY 757 GGTATCGGGGGGCAACAGTCTGAAACTTATTCATGATGATGATGATGATGATGATG 816
    || || || || || || || || || || || || || || || || || || || || || ||
Db 923 GGCATTTGGGCGCAAGCGGCTGTGATGATGATGATGATGATGATGATGATGATGATG 982

```

```

QY 817 TTGGAGATCTTTAATTAAGACAGATATCAATATCTCTAGAGAGCTGGCTTACCAAGAGCT 876
    || || || || || || || || || || || || || || || || || || || || || ||
Db 983 GTGAGCGGCTGCTGACCCAGCAGATACCCGTTCCAGAGAACTGGCTCCACAGAGAGCC 1042
QY 877 CCACGCTTGTTCAGAGAGCTTAATG---TCACATTTGATTTCTCTGAGCTAATATGACT 933
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1043 CAGCAGCTTCTCTGAGAGCAGAACTACTGAGCCAGAGAGTCTGTGAGAGCTGAATGAGAG 1102
QY 934 GCACCTGATGAGAGAGGCTCTCATAGTTTCTCGTGAAGAAATATGTTTCAACAGAT 993
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1103 GAACCAAAAGAAAGAGAGTTGGTCAATTTATGTGTGTGAAGAGATTTTCTGAAGAG 1162
QY 994 CGGCTGACAAAGCCATAGAGAGATCAATCTGCCAAGATTAATGCTGCGCAAGAGAGA 1053
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1163 CGAATTCGAGAGTGGGGTCAAGCGGCTGAGTAAGAGCCGCCAGGCGACAGACCGAGAGCG 1222
QY 1054 CTCGAGTCTTTTTCAGCCAACTGCCACACATGACAGCCGCTAAGAGAGAGACT 1113
    ||||| ||||| ||||| || || || || || || || || || || || || || || || ||
Db 1223 CTCGATGATTTCTTCAGGTGACAGCTCCTACTCTCTAGC-----TAAAGCGC 1270
QY 1114 TCGATTAACAAGCAGAGCAGCTCGAACAAGAAACAAAGCGCTGTGAAAGAGAAA 1173
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1271 AAGAGCCCAAGAACCAAGGCGCTGCTTAAGAAAGAAAGACATGCGGAGACCGGAGAG 1330
QY 1174 T 1174
    |
Db 1331 T 1331

```

RESULT 11
AC026761
LOCUS 166088 bp DNA linear HTG 25-JAN-2002
DEFINITION Mus musculus chromosome 19 clone R22-325122 strain 129/SvEvTacrB,
WORKING DRAFT SEQUENCE, 25 unordered pieces.
AC026761.10 GI:15211712
HTG: HTGS_PHASE1, HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 166088)
Grills,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
Gordon,M., Goltz,J.S., and Kucherlapati,R.
Submitted (24-MAR-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
On Aug 21, 2001 this sequence version replaced g1:10937955.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site:
http://wchanling.bwh.harvard.edu:9088/hpcpg/jsp/hpcpg/sequence/mous
e.html
Contact: gntm@aceped.bwh.harvard.edu
-----Summary Statistics
Center project name: ACD
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100%
*Consensus quality: 156712 at least Q20
*Consensus quality: 152694 at least Q30
*Consensus quality: 147247 at least Q40
Estimated insert size: agarose-FP - N/A
*Estimated insert size: 165608 - sum-of-contigs

Quality coverage: agarose-PF - N/A
Quality coverage: 4 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 34556: contig of 34556 bp in length
* 34557 34576: gap of unknown length
* 34577 62727: contig of 28151 bp in length
* 62728 62747: gap of unknown length
* 62748 81216: contig of 18469 bp in length
* 81217 81236: gap of unknown length
* 81237 92870: contig of 11634 bp in length
* 92871 92890: gap of unknown length
* 92891 103627: contig of 10737 bp in length
* 103628 103647: gap of unknown length
* 103648 114697: contig of 11050 bp in length
* 114698 114717: gap of unknown length
* 114718 123385: contig of 8668 bp in length
* 123386 123405: gap of unknown length
* 123406 131815: contig of 8410 bp in length
* 131816 131835: gap of unknown length
* 131836 132037: contig of 202 bp in length
* 132038 132057: gap of unknown length
* 132058 138473: contig of 6416 bp in length
* 138474 138493: gap of unknown length
* 138494 144023: contig of 5530 bp in length
* 144024 144043: gap of unknown length
* 144044 144924: contig of 881 bp in length
* 144925 144944: gap of unknown length
* 145191 145191: contig of 247 bp in length
* 145192 145211: gap of unknown length
* 145212 146414: contig of 3203 bp in length
* 146415 148434: gap of unknown length
* 148435 151539: contig of 3105 bp in length
* 151540 151559: gap of unknown length
* 151560 154468: contig of 2909 bp in length
* 154469 154488: gap of unknown length
* 154489 156193: contig of 1705 bp in length
* 156194 156213: gap of unknown length
* 156214 157744: contig of 1531 bp in length
* 157745 157764: gap of unknown length
* 157765 159461: contig of 1697 bp in length
* 159462 159481: gap of unknown length
* 159482 160346: contig of 865 bp in length
* 160347 160366: gap of unknown length
* 160367 161824: contig of 1458 bp in length
* 161825 161844: gap of unknown length
* 161845 163663: contig of 1819 bp in length
* 163664 163683: gap of unknown length
* 163684 164031: contig of 348 bp in length
* 164032 164051: gap of unknown length
* 164700 164700: contig of 649 bp in length
* 164701 164720: gap of unknown length
* 164721 166088: contig of 1368 bp in length.

FEATURES

source
1. 166088
/organism="Mus musculus"
/strain="129/SvEvArtBr"
/db_xref="taxon:10090"
/chromosome="19"
/clone="RP2-325J22"
/sex="male"
1. 34556
misc_feature
/note="assembly_name:Contig81"
34577. 62727
/note="assembly_name:Contig80"
misc_feature
62748. 81216

misc_feature /note="assembly_name:Contig79"
81237. 92870
/note="assembly_name:Contig78"
92891. 103627
misc_feature /note="assembly_name:Contig77"
103648. 114697
misc_feature /note="assembly_name:Contig76"
114718. 123385
misc_feature /note="assembly_name:Contig75"
123406. 131815
misc_feature /note="assembly_name:Contig74"
131836. 132037
misc_feature /note="assembly_name:Contig73"
132058. 138473
misc_feature /note="assembly_name:Contig72"
138494. 144023
misc_feature /note="assembly_name:Contig71"
144044. 144924
misc_feature /note="assembly_name:Contig70"
144945. 145191
misc_feature /note="assembly_name:Contig69"
145212. 148414
misc_feature /note="assembly_name:Contig68"
148435. 151539
misc_feature /note="assembly_name:Contig67"
151560. 154468
misc_feature /note="assembly_name:Contig66"
154469. 156193
misc_feature /note="assembly_name:Contig65"
156214. 157744
misc_feature /note="assembly_name:Contig64"
157765. 159461
misc_feature /note="assembly_name:Contig63"
159482. 160346
misc_feature /note="assembly_name:Contig62"
160367. 161824
misc_feature /note="assembly_name:Contig61"
161845. 163663
misc_feature /note="assembly_name:Contig60"
163684. 164031
misc_feature /note="assembly_name:Contig59"
164052. 164700
misc_feature /note="assembly_name:Contig58"
164721. 166088
misc_feature /note="assembly_name:Contig57"
BASE COUNT 44602 a 38872 c 39481 g 42573 t 560 others
ORIGIN

Query Match 22.1%; Score 305.4; DB 2; Length 166088;
Best Local Similarity 56.3%; Pred. No. 4.3e-68;
Matches 642; Conservative 0; Mismatches 481; Indels 18; Gaps 3;

QY 37 ATGGGATTCAGAGCTTTGACGAACACGCTGCGCGGACATATGGCCCAAGCGATGAAGAG 96
DB 11414 ATGGGATTCAGAGCTTTGACGAACACGCTGCGCGGACATATGGCCCAAGCGATGAAGAG 11473
QY 97 CAGAAGTTCAGAGCTTTCGACGAACATCGCGTTCGACGCGCATGAGCATATAC 156
DB 11474 AATGACATCAAGAGCTTTCGACGAACATCGCGTTCGACGCGCATGAGCATATAC 11533
QY 157 CAGTCTGATTTAGTTCGAGAGCAGCATGACCAACTTCACAAATGAAGCTGTGGA 216
DB 11534 CAGTCTGATTTAGTTCGAGAGCAGCATGACCAACTTCACAAATGAAGCTGTGGA 11590
QY 217 GTGACATGATTTTCGAGAGATTTTCACGCGGACATTAAGATTAAGAGCGGATC 276
DB 11591 ACACCGACGACCTGATGCGATGTTTCACCTTACATCCCATGATGGAATGGCATC 11650
QY 277 AAGCAGTTATGTTTTCGAGAGATTTTCACGCGGACATTAAGATTAAGAGCGGATC 336
DB 11651 AAGCAGTTATGTTTTCGAGAGATTTTCACGCGGACATTAAGATTAAGAGCGGATC 11710
QY 337 AGATCTCAAAAAGAGATATGACCAACCAAGATCTGACTGAGGCGATGAGAGTGAAGAT 396

```

Db 11711 CGCAGTGGAGGCGCGCGCGAGGCTGAGAAAGCACTCCAGAGGCTCAGAGGCTGGGATG 11770
QY 397 AAGATGCGATTGAAAAATTGAGCAGAGAGACTGTAAAGGCTACACAAAGCAACACAGAA 456
Db 11771 GAGGAGGAGGTGGAGAAAGTTCCACCAAGAGGCTGTAGAGTACCAACAGCAACATGAT 11830
QY 457 GATTGTAAAGCGCTATTAAAGCTTATGAGGGGTTCTGTTGTAGAGGACCTTTGAAAGCA 516
Db 11831 GAGTGTAAACCTGCTGAGACCTCCTATGGGATCCTTACCTTGAATGACCAAGGAGGCA 11890
QY 517 GAAGCAGATGTCCAGCCCTTTCATPAAAGATAGGTGTGCTGTTGCTTCCAGAAAGAT 576
Db 11891 GAGGCGACCTGTGCTGCGCTGGCAAGAGGCTGGCAAGATGTATGCTGCGGCGACGAGAGAC 11950
QY 577 ATGAGCTCCTTACTTTTGGGGGCTCCAGGTTCTTCTGTCATTTATGATTCAGATTCC 636
Db 11951 ATGAGCTCCTTACTTTTGGGGGCTCCAGGTTCTTCTGTCATTTATGATTCAGATTCC 12010
QY 637 AAGAAATACCTGTGATGAGATTTGATTTGCCAAGTTTGGAGAGCTTGAACCTCAC 696
Db 12011 AAGAGTGCCTCATCAAGAGTTCCATCTGAGCGCGGCTCTGAGAGGCTGGGCTGTAC 12070
QY 697 ATGAGCGCTCATGATTTGATTCCTGATCCTGTGATGATGATTTGATGATGATCAAA 756
Db 12071 CAGGAGCGATTTGATGATCTGTGATCTGTGGGATGAGGACTACTGCGAGAGCATCCGT 12130
QY 757 GGTATCGGGGGGCAAGAGCTGTGAAACTTATTCGTCAACATGAGTGGTCCATAGAAAGCATC 816
Db 12131 GGCATTGGCGCCCAAGCGGCTGTGATCTGATCAGAAACATPAAAGATGAGAGATC 12190
QY 817 TTGGAGATTTTAAATGAGACAGATATCAATTCCTGAGAGCTGGCTTACCAAGAGCT 876
Db 12191 GTGAGCGCGCTGAGACCCAGAGTACCCGTTCCAGAAATCTGGCTCCAAAGAGAGCC 12250
QY 877 CGAGCCTTGTTCAGAGAGCTTAATG--TCACTTGTGATTTCTGAGCTTAATGACT 933
Db 12251 CAGCAGCTTCTCTGAGAGGAGAGAGTACTGAGACCAAGCTGTGAGAGCTGAGAGGAGC 12310
QY 934 GCACATGAGAGAGGCTCATTAAGTTTCTGTGTAAGATTAATGTTTCAAGAGAGAT 993
Db 12311 GAGCCAAATGAAAGAGAGTTGGTCAATTTATGTGTGTAAGAGACATTTTTCGAAGAG 12370
QY 994 CGGAGTGAAGAGGCTATGAGAAATCAATCTGCCAACAATAATCTGTCGAAGAGAGA 1053
Db 12371 GCAATTCGCACTGGGCTCAAGCGCTGAGTAAAGAGCGCGCAGAGGAGCAAGAGAGAGC 12430
QY 1054 CTGCACTCCTTTTCAAGCCCACTGCCACCATCAGACCGCTTAAAGAGAGAGACT 1113
Db 12431 CTGCACTCCTTTTCAAGCGCTGAGAGGCTCCTCTCTCAGC-----TAAAGCGC 12478
QY 1114 TCGGATTAAGCAAGCAGGAGGCTGCGAACAAGAAACAAAGGCTGTGTAAGAGAGAAA 1173
Db 12479 AAGAGCCCAAGCAAGGAGGCGCTGCTAAGAGAAAGCAAGAGCTGGGAGAGCGGAGAG 12538
QY 1174 T 1174
Db 12539 T 12539

```

```

RESULT 12
LOCUS      HSRAD2      1757 bp      mRNA      linear      PRI 07-DEC-1998
DEFINITION Homo sapiens mRNA for flap endonuclease-1.
ACCESSION  X76771
VERSION    X76771.1 GI:1905802
KEYWORDS   fe1 gene; flap endonuclease-1.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1757)
AUTHORS   Watts,F.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (16-DEC-1993) F. Watts, University of Sussex, School of
REFERENCE  Biological Sciences, Falmer, Brighton BN1 9QG, UK
AUTHORS    Murray,J.M., Mayasoli,M., al-Hatithy,R., Sheldrick,K.S.,
            Lehmann,A.R., Carr,A.M. and Watts,F.Z.
TITLE      Structural and functional conservation of the human homolog of the
            Schizosaccharomyces pombe rad2 gene, which is required for
            chromosome segregation and recovery from DNA damage
JOURNAL    Mol. Cell. Biol. 14 (7), 4878-4888 (1994)
MEDLINE    94277093
FEATURES   location/Qualifiers
            1..1757
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /gene="fen1"
            /gene="fen1"
            /gene="fen1"
            /product="flap endonuclease-1"
            /protein_id="CA54166.1"
            /db_xref="GI:3980293"
            /translation="MGIOGLAKLNDVAPSLIRENDIKSYFGRKVAIDMSITYPLI
            AVRGSDVLONEGETSHLMGFRTIRRMENIKPYVFDKPKQLSGLARKE
            RRAEARKQLQQAAGAEDEVEKFTKRLVYVQHNDECKHLISLNGTILDAPEAE
            ASCALVLRKAGKYAAATDEMDCLTGSPVLMRLHLSSEAKKLPIDFHLSTRLOELG
            NQOFLVDLICILGSDYCSIRIGIPKRAVDLIDKHSIEIVRLDPNPKYVPVENMLH
            KEAHOLFLEPEVLDESEVLEKMPSENEELIKFMCGRKQSESRIRISRLSKSROG
            STQGRIDPEFVKYGLSSAKRKEPEPGRSTRKKRKTGAACKFPRK"
BASE COUNT 418 a 443 c 542 g 354 t
ORIGIN
Query Match      22.0%; Score 304.2; DB 9; Length 1757;
Best Local Similarity 54.6%; Pred. No. 4e-68;
Matches 675; Conservative 0; Mismatches 553; Indels 9; Gaps 3;
QY 2 GACCAGCGCTCGCGCCACAGCCGCGCAGAGAGATGGGCAAGGTTTGAAGAAC 61
Db 333 GTCCCAAGGCGCATCATCCCTCTCTGTGTTGCCATGGAATTCAGAGCTGGCCAAAC 392
QY 62 TCGTGGCGGACATGCGCCCAAGCGATGAGAGAGCAAGTGGAGAGACTTGGCGC 121
Db 333 TAATTGCTGATGTGGCCCCCAGTGCCATCCGGAGATGACATCAAGAGCTACTTGGCC 452
QY 122 GCAAAATCGCCGTCGACGCCAGCATGAGCATATACCACTTCTGATTTGTAGTGAAGA 181
Db 453 GTAAGTGGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511
QY 182 CAGCATGGAACCTCCCAAAATGAAGAGTGGAGTCACTAGTCAATTTGCAAGAGATGT 241
Db 512 --GGTGGGATGTGCTGCAAGATGAGAGAGGTGACACCAAGCCACCTGTGAGGCAATG 569
QY 242 TCAACCGGCAATTAAGATTACTGGAAGCGGAGATCAAGCCAGTTTATGTTTGTAGTGA 301
Db 570 TCTACCGCACCATTCGATGATGATGAGAACGCGCATCAACCGGTGATGTGATGATGGA 629
QY 302 AGCCTCTGATTAATGAAGAAACAGAGCTTGTCTAAAGATTAATCAAAAAGAGATGATGCA 361
Db 630 AGCGCCACAGCTCAATGATGAGAGAGCTGCGCAACGCAATGAGCGGCGGCTGAGAGAG 689
QY 362 CCAAGATCTGACTGAGGAGCTAGAGTGAAGATGAAGATGCGATTTGAATAATTGAGCA 421
Db 690 AAGAGCGCTGACAGACAGGCTCAGGCTGCTGGGCGGACAGAGAGTGAATAAATTCCTA 749
QY 422 AGAGAGCTGTAAGAGTCAACAGCAGACAGAGAGATTTGAAGCGGCTTTTGAAGACTTA 481
Db 750 AGGCGTGTGAAGGTGATCAACAGCAATGATGATGATGATGATGATGATGATGATGATG 809
QY 482 TGGGAGTTCCTGTTTGAAGGACCTTTGAAGCAAGAGCAATGTGACCCCTTTGCA 541
Db 810 TGGGATTCCTTTATTTGATGACCCAGAGTGAAGAGAGGCGCAGCTGTGCTGCTGTGA 869

```

```

QY 542 TAAACGTAAGGTCGCTGCTTCAGAAAGATATGACTCCCTTACTTTTGGGCTC 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 870 AGGCTGGCAAGATCTATGCTGCGTACCGAGACATGACTGCTCACCCTGGCGCC 929
QY 602 CAGGTTCTCTGTCATTTAATGATTCACAGTTCACAGAAATACCTGTGATGAAATTTG 661
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 930 CTGTGCTAATGGACACCTGACTGTCAGTGAAGCCAAAAGCTGCATTCAGGAATTC 989
QY 662 ATGTGGCAAGTTTGGAGAGCTTGAATCTACCATGAGCACTTATTTATTTGCA 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 990 ACCGTGACCGGATTTGCGAGCTGGGCTGAACACGAGAACTGTGTGATGATGCA 1049
QY 722 TCTGTGTGATGATGATTTGTGATAGCATCAAGATATCGGGGGGCAACAGCTTGA 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1050 TCTGTGAGCAAGTACTACTGTGAGAGATTCGGGGGATTTGGGCCCAAGCGGCTGTG 1109
QY 782 AACTTATTTGTCACATGGTTCATAGAAAGATCTGGAATCTTATTAAGACAT 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1110 ACTCATCCGAGACACAGAGATCGAGAGATGCTGCGGACCTTGACCCCAACAGT 1169
QY 842 ATCAAAATTCGTAGAGCTGCTTACCAAGAGCTCGAGCTTGTTCAGAGAGCTATG 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1170 ACCCTGTGCGCAAAATTTGCTTCACAAAGAGGCTCACAGCTCTTCTTGAACCTGAG 1229
QY 902 ---TCACATTTGATTTCTCTGAGCTAAATGAGTGCACCTGATGAGGAGGCTCATTA 958
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1230 TGTGTGACCCAGAGCTGTGAGAGCTGAAGTGAAGCAGCAATGAAAGAGCTGATCA 1289
QY 959 GTTTCCTGTGTAAGAAATGTTTTCACAGAGATCGGGGTGCAAAAGCCATAGAGA 1018
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1290 AGTTCATGTTGTGTAAGAGAGTTCCTGAGAGACGAAATCGGAGGTGGGTCAAGAGC 1349
QY 1019 TCAATCTGCGCAAGATTAATGCTGCGCAAGAGAGCTGAGCTTTTTCAGGCACTG 1078
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1350 TGAGTAAAGACCCCGCAAGGACAGACCCAGCGGCTGATGATTTCTTCAAGGTACCG 1409
QY 1079 CCAC---CACATGACGACCGCTAAAGAGAGAGACTTGGATTAAGCAAGCAGCAG 1135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1410 GCTCAGCTCTTCTGAGTAAGCGGAGAGCCAGAAAGCCAAAGGATCCATTAAGAGAG 1469
QY 1136 CTCGCAACAGAAAGAGAGCTGTGTAAGAAAGAAATATCTGATGCTTATGATAC 1195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1470 CAAGAGCTGGGGCAGCAGGAGGATTAAGAGGGAGAAATATGCTTCCCATTTATC 1529
QY 1196 AACTAGCACTACGAAGCAGCGGTGCGCTGATCACTT 1232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1530 CTCCTTCCCGCAGAAATATTGGCTTGTACCTT 1566

RESULT 13
BC000323          2031 bp      mRNA      linear      PRI 12-JUL-2001
LOCUS             Homo sapiens, flap structure-specific endonuclease 1, clone
DEFINITION        MGC:8478 IMAGE:2821792, mRNA, complete cds.
ACCESSION         BC000323
VERSION           BC000323.1 GI:12653112
KEYWORDS          MGC.
SOURCE            human.
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2031)
Strausberg, R.
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
REMARK
COMMENT

```

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNT)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nih.gov
Shechenko, Y., Welherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-T., Karlins, E., Legaspi, R.,
Lim, M., Madino, O.L., Mastello, C., Mastriani, S.D., McCloskey, J.C.,
McGowan, J., Pearson, R., Snyder, B., Stancil, P., Thomas, P.J.,
Tongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNT at: http://lmage.llnl.gov
Series: IRAL Plate: 1 Row: 9 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 1905802.

FEATURES
    source
        1..2031
        /organism="Homo sapiens"
        /db_xref="LocusID:2237"
        /clone="MGC:8478 IMAGE:2821792"
        /tissue_type="lung, small cell carcinoma"
        /clone_lib="NIH-MGC-7"
        /lab_host="DH10B-R"
        /note="Vector: pOTB"
        139..1281
        /product="flap structure-specific endonuclease 1"
        /protein_id="AAH00323.1"
        /db_xref="gi:12653113"
        /translation="MGI:GALADLADYAPSAIIBNDIKSYRKAVIDASISYOPILI
        AYRGGSDVLONEGETTSHLMGFEYRTIMNGIKPYVDPGKRPOLKSELIKRSB
        RRAAEKQDLOQAGABDEVEKTRKLVKTRKONHDECKHLISLGLPIYDASBEB
        ASCEALVAKRIVTAATVEDMDCTFESPVLMHLSAEKRLPIDEHPLSTLDELGL
        NQEOEVLICLISDYCESLRIGIPRAVDLQKRSLEIVRLDPKRYVPPNWLH
        KEAHOFLFEPEVDELSVAKRKEPEKESERIRTSVSKRLSKSRQ
        STQORLDFEFTVGLSLSAKRKEPEKESKRRKATGAAGKFRKRGK"

BASE COUNT      533 a      465 c      589 g      444 t
ORIGIN
Query Match      22.0%; Score 304.2; DP 9; Length 2031;
Best Local Similarity 54.6%; Pred. No. 4.1e-68;
Matches 675; Conservative 0; Mismatches 553; Indels 9; Gaps 3;
QY 2 GACCCAGCGTCCGCGCAGCCGCCGAGACGAGATGGGCATCAAGGTTTGCAGAAC 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 104 GTCCCAAGGGCGAGTCATCCCTCCTGCTGTTGCCATGGGAATTCAGAGCCGCAAC 163
QY 62 TGTGTGCGGACATATCGCCCGCAAGCGATGAAGAGCAGAACTTGCAGAGCTACTTGGCC 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 TAATGCTGATGTGCCCCCGACGTCCATCCGGAATGACATCAAGACTACTTGTGGCC 223
QY 122 GCAAAATCGCGCTCCAGCCGACGATGACATATACAGTTCTGTGATTTGTGGAAGGA 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 GTAAGGTGGCCATGATGATCCCTCTATGACCATTTATTCAGTTCTGATTTGCTGGCCAG 282
QY 182 CAGCGATGGAATCTCCAAATGAAGTGTGTAAGTCACTAGTCACTTGCAGGAATGT 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 ---GGTGGGATGTGCTGCAAGATGAGAGGAGTACACACGACGACCTGATGAGCATGT 340
QY 242 TCACCGGCAATTAAGATTACTGGAAGCGGAATCAAGCAGTTATGTTTATGATGAGA 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 TCTACCGCACCACTTCGATGATGAGAGACGACATCAAGCCCGGTATGTTGATGAGCA 400
QY 302 AGCTTCCTGATTAAGAGAAACAGAGCTTGTCTAAAGTACTCTAAAGAGATGATGACA 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 AGCGGCGACAGCTCAAGTCAGCGGAGCTGGCCAAACGATGAGCGGCGGTGAGAGCAG 460
QY 362 CCAAAGATCTGACTAGGCGAGTAGAGGTAGAGATTAAGATGCGATTGAAGAAATTGAGCA 421

```

Accession	Sequence	Position
Db	AGAAAGCAGCTGCACGACAGCTCTAGGCTGTGGGGCCGAGCAGAGAGGTGGAAAAATTCTACTA	520
Oy	AGAGAGACTGTAAAGGTCACCAAGGCACAGCACAAGATTGTAACGGCTATTAAAGACTTA	481
Db	AGGGCGTGGTGAAGGTCACCTAAGCAGCACAATGTGATGCAAAACATCTGGCTGAGGCTCA	580
Oy	TGGGGGTTCTGTGTGTAAGAGCACCTTCTGAAGCAGACAGCAATGTGCAAGCCTTTTGCA	541
Db	TGGGCATCTCCCTTATCTTGTATGACCCAGTAGAGCAGGCGAGCTGCTCCCTGGTGA	640
Oy	TAAACGATTAAGGTTTCGCTGTTCCTCAGAGATATGAGACCTCTACTTTTGGGGCTC	601
Db	AGGCTGCAAAAGCTTATGTGTGGGGCTACGAGAGCAGATGAGATGCTCTCACCTTTGGCAGCC	700

```
misc_feature      1. .20726
                  /note="assembly_fragment"
misc_feature      20827. .34666
                  /note="assembly_fragment"
misc_feature      34767. .46065
                  /note="assembly_fragment"
```

Query Match	21.98;	Score 302.8;	DB 2;	Length 138627;
Best Local Similarity	55.08;	Pred. No. 2e-67;		
Matches 661; Conservative	0;	Mismatches 532;	Indels 9;	Gaps 3;

QY	37	ATGGCATCAAGGCTTTGACGAAACTGCTGGCGGACCAATGGCCCAAGCGCATGACAGAG	96
Db	17322	ATGGGAATTTAAAGCCCTGGCGCAAACTAATGTGTGATGGGCCCCCATGTGCATTCGGGAG	17381
QY	97	CAGAAGTTTCAGAGACTACTTCGCGCGCAAAATCGCCGTCGACGGCCAGCATGAGCATATAC	156
Db	17382	AATGACATTCAGAGACTACTTTGGCCGTAAGGTCGACATTCGATTCGCTATGAGCATTTAT	17441
QY	157	CAGTTCCTGATTTAGTGTGAAGAGCAGCATGACAACTCCAAATTAAGTCGTGGCA	216
Db	17442	CAGTTCTGATTTCTGTGTGGCAG --- GGTGGGATTTGCTGTGCAATTAAGAGAGGTGAG	17498
QY	217	GTCACTAGTCAATTTGCAAGAAATGTTCAACCGGACATAGATTTACTGGAACGGGCATC	276
Db	17499	ACCAACAGCCACTGATGGCATGTTCCTACCGCACCATTCGCAATGAAGGAGAAAGGCATC	17558
QY	277	AAGCAGTTTATTTTGTGATGTGCAAGCCTCCTGATATGAAGAAACAAGACTTGTCTAA	336
Db	17559	AAGCCGCTATGTCTTTGATGTGCAAGCGCCACAGCTCAAGTCAGCGCAGATGTGGCANA	17618
QY	337	AGATATCAAAAAAGATGATGTGAACCAAAAGTTCGACTGAGGCGATGAGATGATGAGAT	396
Db	17619	CGCATGTAGCGCGGGGTGAGGAGAAAGCAAGCTGTGACAGCGGTCTCAGGCTGTGGGGCC	17678
QY	397	AAACATCGAATGAAAAATTGAGCAAGAGACTGTAAAGGTCCACAAGGCAACACAACGAA	456

```

Db 17679 GAGCAGAGAGTGGAAAAATTTCTACTAGCGGCTGGTAGAGTCACTAGACAGCATATAT 17738
QY 457 GATTGTAACGGCTATTATTAAGCTTATGGGGTTCCTGTTGTAGAGGACCTTGTGAAGA 516
Db 17739 GAGTGAACAATCTGCTGAGACCTCATGAGCATCCCTTATCTTATGATGACCCAGTGA 17798
QY 517 GAGGCAACAATGTCAGCCCTTTGCTCAATAACGTAAGTGTGGCGCTTGTGCTTCAACAAT 576
Db 17799 GAGGCAAGCTGTGCTGCTGCTGAGAGCTGGCAAGATCTATGCTGCGGCTACCGAGAC 17858
QY 577 ATGAGACTCCCTTACTTATTTGGGGCTCCAGCGTTCCCTGTCATTTAATGATCCAAATT 636
Db 17859 ATGAGACTCCCTTACTTATTTGGGGCTCCAGCGTTCCCTGTCATTTAATGATCCAAATT 17918
QY 637 AAGAAATACCTGATGATGATTTGATGTTGCAAGCTTTTGGAGAGCTTGAACCTAC 696
Db 17919 AAAAAAGCTGCAATCCAGAGATTCACCTGAGCGGATTTCTCAGAGAGCTGGGCTGAC 17978
QY 697 ATGAGACAGTTTCAATTTGTCATCTGTTGATGATGATGATGATGATGATGATGATGAT 756
Db 17979 CAGGACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18038
QY 757 GATATGCGGGGGGCAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
Db 18039 GATATGCGGGGGGCAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 18098
QY 817 TTGAGATCTTATTAATAAGACAGATATCAAAATTCAGAGAGCTGGCTTACCAAGAGCT 876
Db 18099 GTGGGCGGCTTGAACCCCAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 18158
QY 877 CGAGCTGTTGCAAGAGCCTATATG---TCACATTTGATATTCCTGAGCTTAATGAGCT 933
Db 18159 CACCAAGCTCTTCTTGGAGCTGAGTGTGCTGAGCCAGAGCTCTGTGAGAGTGAAGT 18218
QY 934 GCACCTGATGAGAGGCTCTCATAGTTTCTGTGTAAGATATGTTTCAACAGAGAT 993
Db 18219 GAGCAATATGAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 18278
QY 994 CGGCTGCAAGAGGCAATAGAGATCAAAATTCGCCAAGATTAATTCGTCGCAAGAGA 1053
Db 18279 CGAATCCGCGAGTGGGCTGAGAGAGCTGATGATGATGATGATGATGATGATGATGAT 18338
QY 1054 CTGAGTCTCTTTTCAAGCCACTGCCAC---CACATCAGCAACCGCTAAACGAGAGAG 1110
Db 18339 CTGATATATTTCTTCAAGGAGCCGCTCACTCTTCTTCAAGGAGGAGGAGCCAGAA 18398
QY 1111 ACTGCGATTAACCAACAGCAGCGAGCTGGGAAACAAAGAGCTGTGGAAGAG 1170
Db 18399 CCCAAGGAGATCCACTAAGAAAGAGCAAGAGCTGGGAGCAGAGAGTTTAAGAGGGA 18458
QY 1171 AATATATCTTGGATGCTTGTATGATCAACTAGACATGAGAGAGCGGTGCGTGTATC 1230
Db 18459 AATATATCTTGGATGCTTGTATGATCAACTAGACATGAGAGAGCGGTGCGTGTATC 18518
QY 1231 TT 1232
Db 18519 TT 18520

```

```

RESULT 15
AP002380/c 169053 bp DNA linear HTG 31-MAY-2000
LOCUS Homo sapiens chromosome 11 clone RP11-467L20 map 11q, WORKING DRAFT
DEFINITION SPOUNCE: 35 unordered pieces.
ACCESSION AP002380
VERSION AP002380.1 GI:8131644
SOURCE HTG: HTGS_PHASE1; HTGS_DRAFT.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 169053)

```

AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. Homo sapiens 169,053 genomic DNA of 11q Published Only in Database (2000) In press 2 (bases 1 to 169053)

Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. Direct Submission Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8535, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923, Fax: 81-42-778-9924)

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDrafl1
Center clone name: RP11-467L20
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 147827 bases at least Q40
Consensus quality: 157936 bases at least Q30
Consensus quality: 162410 bases at least Q20
Insert size: 165653; sum-of-contigs
Quality coverage: 4.43x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and the order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1	21322	contig of	21322	bp in length
21423	40245	contig of	18823	bp in length
40346	50492	contig of	10147	bp in length
50393	58569	contig of	7977	bp in length
58670	67385	contig of	8716	bp in length
67486	75248	contig of	7763	bp in length
75349	84453	contig of	9105	bp in length
84554	89492	contig of	4939	bp in length
89593	96445	contig of	6833	bp in length
96546	102631	contig of	6066	bp in length
102732	107199	contig of	4468	bp in length
107300	111820	contig of	4521	bp in length
111921	117068	contig of	5148	bp in length
117169	120901	contig of	3733	bp in length
121002	124072	contig of	3071	bp in length
124173	127611	contig of	3439	bp in length
127712	129401	contig of	1690	bp in length
129502	133141	contig of	3640	bp in length
133242	136380	contig of	3139	bp in length
136481	139425	contig of	2945	bp in length
139526	143281	contig of	3756	bp in length
143382	146449	contig of	3068	bp in length
146550	148828	contig of	2279	bp in length
148929	151045	contig of	2217	bp in length
151146	153372	contig of	2227	bp in length
151473	154996	contig of	1554	bp in length
155097	156763	contig of	1667	bp in length
156864	159132	contig of	2269	bp in length
159233	161349	contig of	2117	bp in length
161450	162850	contig of	1401	bp in length
162951	164533	contig of	1583	bp in length
164634	165829	contig of	1166	bp in length
165930	167130	contig of	1201	bp in length
167231	168290	contig of	1060	bp in length

Tue May 27 16:02:20 2003

us-09-805-311-5.std.rge

Page 18

Db	3399	GAGCAGAGGATGGA	AAATTCACTAAAGCGGCTGTGAAGGTCAC	TAAACACCAATGAT	3340
Oy	457	GATTGTAAAGCGTATT	TAAGACTTAAGGGGGTTCGTGTGAAGGCACTTC	GAAGCA	516
Db	3339	GAGGCAAAATCTGGCT	GAGACCTCAAGGGCATCCCTTATCTTGATGCA	CCACAGTGAAGCA	3280
Oy	517	GAACGACAATGTGC	ACCCCTTTGCATAAACGATTAAGGTTCGCTGTG	CTTCAGAAAGT	576
Db	3279	GAGCGCAGCTGTCT	CTCCCTCGGTGAAGGCTGGCAAAAGTCATAGCTCGCGCTACCGAGAGAC	3220	
Oy	577	ATGAGACCCCTTACT	TTTGGAGCTCCACGGTTCCTTCGTCAATTAATGATCA	CAAGTTCC	636
Db	3219	ATGAGACTGCTCAC	CTTCGCGACCCCTGTGTCTAATGCGACACCTGAC	TGCGACATGAAGCC	3160
Oy	637	AAGAAATACCTGT	GATGGAATTTGATGTTGCCAAGGTTTGGAGACCTTGA	CACTCAC	696
Db	3159	AAAAAGCTGC	CAATCCAGGAATTCACACCTGACCCGGATTCCTGACGAGACTGGGCTGAAC	3100	
Oy	697	ATGACACAGTTCAT	TGATTTGTGTGCATCCCTGTGTGATGTGACTATTTGAT	TGAAGATCAAA	756
Db	3099	CAGGAAGATTTGG	TGTGATCTGTGCATCTCTGTAGGACGTAGCATCTG	GAAGATATCGG	3040
Oy	757	GCTATCGGGGGGCA	ACAGCTGTGAACCTTATTCGCACATGGGCTCATG	AAAGATC	816
Db	3039	GGTATTTGGGCC	CAAGCGGGCTGTGACCTTCATCCAGAACACAAAGATGAGAGACAT	2980	
Oy	817	TTTGAGAATCTTA	TATAAAGACAGATATCAAAATTCCTGAGGACTGGCTTAC	CAAGAACT	876
Db	2979	GTTGGGAGCATTA	CAACCCCAAGTACCTGTGCGAGAAATTTGGCTCCACAGAGAGCT	2920	
Oy	877	CGAGCGTTTGTCA	AGGAGGAGCTTAATG---TCACATTTGATTAATCC	TGAGCTAAATGACT	933
Db	2919	CACAGCTCTTCT	TGGAACCTGTAGTGCTGGAACCAAGATCTGTGTG	AGACTGAAGTGAGC	2860
Oy	934	GCACCTGATGAG	AGGCTCTCATTAAGTTTCTGTGTAAAGATTAATGTG	TTTCAACGAAAT	993
Db	2859	GAGCCAATGA	AGAGGCTGATCAAGTTCAATGTGTGTGA	AAACCAAGTCTCTGAAGAG	2800
Oy	994	CGGGTGA	CAAAAGCCATAGAAAGATCAAAATCTGCCAA	GAATTAATCTGTGCAAGAGA	1053
Db	2799	CGAATCCG	CACTGTGGGTCAGAGGCTGAGTAAGAGCCGCGCAAGCAGACAC	CCAGGGCGC	2740
Oy	1054	CTCAGATCCCTT	TTTCAAGCCCACTGGCAC---CACATTCAGACCCGCT	TAAACGGAGAG	1140
Db	2739	CTGGATGATTTCT	TCAAGTGACGCGGTCACTCTCTTCAAGCTAAACGCA	AGGACAA	2680
Oy	1111	ACTTGGATTA	AACACAGCAGGACGCTCGCAGACAGAA	AAACAAAGGCTGTGTGAAGAAG	1170
Db	2679	CCCAAGGATTC	CACTAAGAAAGGCAAAACACTGGGGCAGCAGGGA	AACTTTAAAGGGGA	2620
Oy	1171	AAATTAATCTTG	ATGATCTTGATTAACACTAGCACTCGA	AAACACAGCGGTGTGATAC	1230
Db	2619	AAATTAATGTGT	TTCGCCCATTAATACCTCTTCAACCCCA	GAATATTTGGCGTCTGTAC	2560
Oy	1231	TT	1232		
Db	2559	TT	2558		

Search completed: November 5, 2002, 12:54:45
Job time : 2309.69 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

Om nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:10:48 ; Search time 164.876 Seconds
(without alignments)
14380.822 Million cell updates/second

Title: US-09-805-311-5
Perfect score: 1381
Sequence: 1 cgaccacagcgtccgcccac.....ttgaaaaaaaaaaaaaa 1381

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
```

Database :

1	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1980.DAT.*
2	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1981.DAT.*
3	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1982.DAT.*
4	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1983.DAT.*
5	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1984.DAT.*
6	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1985.DAT.*
7	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1986.DAT.*
8	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1987.DAT.*
9	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1988.DAT.*
10	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1989.DAT.*
11	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1990.DAT.*
12	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1991.DAT.*
13	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1992.DAT.*
14	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1993.DAT.*
15	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1994.DAT.*
16	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1995.DAT.*
17	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1996.DAT.*
18	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1997.DAT.*
19	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1998.DAT.*
20	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA1999.DAT.*
21	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA2000.DAT.*
22	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA2001.DAT.*
23	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA2002.DAT.*
24	/SIDS1/gcgdata./geneseq./geneseqn.-emb1./NA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1381	100.0	1381	21	AAAZ7925	Maize Rad2/FEN-1
2	1350.2	97.8	1463	21	AAAZ7923	Maize Rad2/FEN-1
3	1348.6	97.7	1541	21	AAAZ7924	Maize Rad2/FEN-1
4	1334.2	96.6	1478	21	AAAZ7926	Maize Rad2/FEN-1
5	302.6	21.9	1144	20	AAAX02117	Human FEN-1 DNA
6	302.6	21.9	1144	20	AAAX02117	Human FEN-1 DNA
7	276.6	20.0	2033	20	AAAX02111	Human FEN-1
8	276	20.0	1300	23	ABLI4287	Human FEN-1
9	272.2	19.7	1930	20	AAAX02108	Drosophila melanogaster FEN-1 CDNA

10	270.2	19.6	1149	20	AAJ021019
11	252.6	18.3	3362	23	ABJ142866
12	173.2	12.5	5471	23	ABJ100944
13	153.6	11.6	1023	18	AAV766685
14	159.6	11.6	1023	19	AAV598400
15	159.6	11.6	1023	19	AAV33951
16	152.8	11.1	1054	19	AAV33964
17	149.6	10.8	1115	19	AAV339922
C 18	141	10.2	510	22	AAK37479
C 19	141	10.2	510	22	AAJ182866
C 20	135.6	9.8	1032	20	AAJ18560
C 21	134	9.2	349980	22	AAH112240
C 22	128	9.3	514	19	AAV33960
C 23	126.4	9.2	1164	19	AAV33975
C 24	125.6	9.1	1115	19	AAV33978
C 25	119	8.6	1729	23	AAJ33732
C 26	110.2	8.0	889	19	AAV33974
C 27	98.6	7.1	1011	19	AAV65913
C 28	98.6	7.1	1011	19	AAV40009
C 29	97.6	7.1	1164	19	AAV33969
C 30	94.8	6.9	386	19	AAV33979
C 31	90	6.5	546	22	ABJ33065
C 32	90	6.5	546	22	AAK11400
C 33	90	6.5	546	22	AAK37265
C 34	90	6.5	546	22	AAI18098
C 35	89.2	6.5	981	18	AAI76682
C 36	89.2	6.5	981	19	AAV33948
C 37	86.6	6.3	245	22	AAK50354
C 38	86.6	6.3	245	22	AAI7416
C 39	86	6.2	981	19	AAV5837
C 40	76.2	5.5	777	19	AAV33977
C 41	74.6	5.4	840	19	AAV40101
C 42	69.4	5.0	987	19	AAV33976
C 43	67.8	4.9	296	19	AAV33965
C 44	62.8	4.5	514	19	AAV33965
C 45	62.4	4.5	4351	21	AAA7012

ALIGNMENTS

RESULT 1

ID	AAA27925	standard; cDNA; 1381 BP.
...		

AC AAA27925;

DT 12-SEP-2000 (first entry)

DE Maize Rad2/FEN-1 cDNA.

KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;

KW endonuclease; exonuclease; DNA repair; gene targeting; ssDNA

zee may.

key	ET
key	ET

XXI

W0200036109-AL-
FN
XX

22-JUN-2000.
PD
XX

PF 10-NOV-1999; 99WO-052/14/

13-DEC-1990, 2003-0112032Z
FM
XX

(F L O N) F L O N E R I I B N E D I N I I N C
F F
X X

XX
Maria Jall F.B.

DN MIT, 2000 402020/22

DR p-PSDB: AAY95309.

XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants.

PS Example 1: Page 75-76; 85bp; English.

XX The present sequence is that of maize cDNA coding for RAD2/FEN-1
CC (see AAY95309). The cDNA was isolated from a cDNA library prepared
CC from W23 line tassel (16-18 cm long) polyA RNA. Rad2/FEN-1 is a
CC structure specific endonuclease which under certain conditions also
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.

XX Sequence 1381 BP; 441 A; 269 C; 346 G; 325 T; 0 other:

Query Match 100.0%; Score 1381; DB 21; Length 1381;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCCACGGGTCCGGCCACAGCCGCCGACAGAGATGGGATCAAGGTTGACGAAA 60
DB 1 CGACCCACGGGTCCGGCCACAGCCGCCGACAGAGATGGGATCAAGGTTGACGAAA 60
QY 61 CTGCTGGCGGACATGCGCCCAAGGCGATGAAGAGCAGAAGTTGAGAGCTACTTCGGC 120
DB 61 CTGCTGGCGGACATGCGCCCAAGGCGATGAAGAGCAGAAGTTGAGAGCTACTTCGGC 120
QY 121 CGCAAAATCGCCGTCGACGCGACAGATGATGATACCACTTCTGATTTGATTTG 180
DB 121 CGCAAAATCGCCGTCGACGCGACAGATGATGATACCACTTCTGATTTGATTTG 180
QY 181 ACAGGATGGAATCTCACAAATGAAGTGTGAAGTCACTAGTCACTTTGCAAGATG 240
DB 181 ACAGGATGGAATCTCACAAATGAAGTGTGAAGTCACTAGTCACTTTGCAAGATG 240
QY 241 TTCAACCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
DB 241 TTCAACCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
QY 241 TTCAACCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
DB 241 TTCAACCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
QY 301 AAGCTCTCTGATATGAAGAAACAAGAGCTTCTAAAGATATCTCAAAAAGATGATCA 360
DB 301 AAGCTCTCTGATATGAAGAAACAAGAGCTTCTAAAGATATCTCAAAAAGATGATCA 360
QY 361 ACCAAAGATCTGACTGAGGCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 420
DB 361 ACCAAAGATCTGACTGAGGCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 420
QY 421 AAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 480
DB 421 AAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 480
QY 481 ATGGGGGTTCTCTGTTAGAGGCACTTCTGAAGCAAGAGATGTGCACCTTTGCG 540
DB 481 ATGGGGGTTCTCTGTTAGAGGCACTTCTGAAGCAAGAGATGTGCACCTTTGCG 540
QY 541 ATAAACGATTAAGTGTGCTGCTTGAAGATATGATGATGATGATGATGATGATGATG 600
DB 541 ATAAACGATTAAGTGTGCTGCTTGAAGATATGATGATGATGATGATGATGATGATG 600

QY 601 CCACGGTTCCTTCGTCATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 CCACGGTTCCTTCGTCATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTGACATGACATGACATGATGATGATGATG 720
DB 661 GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTGACATGACATGACATGATGATGATGATG 720
QY 721 ATCCGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 ATCCGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 AAATTTATTCGTCACATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 AAATTTATTCGTCACATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 TATCAATTCCTGAGGAGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 TATCAATTCCTGAGGAGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 GTACATTTGATATTCCTGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 901 GTACATTTGATATTCCTGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 TTCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 TTCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 AAATTCGCCAAGATTAATTCGTCGACGAGAGAGCTGATGATGATGATGATGATGATGATG 1080
DB 1021 AAATTCGCCAAGATTAATTCGTCGACGAGAGAGCTGATGATGATGATGATGATGATGATG 1080
QY 1081 ACCACATTCAGCAGCTGTAATGAGAGAGAGCTGATGATGATGATGATGATGATGATGATG 1140
DB 1081 ACCACATTCAGCAGCTGTAATGAGAGAGAGCTGATGATGATGATGATGATGATGATGATG 1140
QY 1141 AACCAAGAAACAAGGCTGTGAGAGAGAGAGATTAATCTTGATGCTTGAATGATGATGAT 1200
DB 1141 AACCAAGAAACAAGGCTGTGAGAGAGAGAGATTAATCTTGATGCTTGAATGATGATGAT 1200
QY 1201 CGACTGAGAAAGCAGCGGTGGGAGTACATGCTGATGATGATGATGATGATGATGATGATG 1260
DB 1201 CGACTGAGAAAGCAGCGGTGGGAGTACATGCTGATGATGATGATGATGATGATGATGATG 1260
QY 1261 ACTCAGAGCTTTGTTAAAGTTCGCCATGTTTCAAGCTGGGTTAAGTTGATGATGATGATG 1320
DB 1261 ACTCAGAGCTTTGTTAAAGTTCGCCATGTTTCAAGCTGGGTTAAGTTGATGATGATGATG 1320
QY 1321 AAGAGATTTGTTACAGTAAACAATTAATGCTGTTTGTGATGATGATGATGATGATGATG 1380
DB 1321 AAGAGATTTGTTACAGTAAACAATTAATGCTGTTTGTGATGATGATGATGATGATGATG 1380
QY 1381 A 1381
DB 1381 A 1381

RESULT 2
AAAZ7923
ID AAAZ7923 standard: cDNA; 1463 BP.
XX
AC AAAZ7923;
XX
DT 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 cDNA.
XX
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
XX endonuclease; exonuclease; DNA repair; gene targeting; ss.
OS
XX
FH key Location/Qualifiers

FT CDS 85..1224
FT /*tag= a
PM WO200036109-A1.
XX 22-JUN-2000.
XX 16-NOV-1999; 99WO-US27147.
XX 15-DEC-1998; 98US-0112332.
XX (PION-) PIONEER HI-BRED INT INC.
XX Mahajan PB;
XX WPI: 2000-452026/39.
XX P-PSDB: AAY95307.
XX
XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
XX recombination and repair in transgenic plants, e.g. for gene targeting
XX and the production of male sterile plants -
XX
XX Claim 1; Page 69-71; 85pp; English.
XX
XX The present sequence is that of maize cDNA coding for RAD2/FEN-1
XX (see AAY95307). The corresponding RNA was isolated from immature
XX ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a
XX structure specific endonuclease which under certain conditions also
XX acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
XX produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
XX The protein is involved in the regulation of DNA repair and
XX recombination in plant systems and therefore may be used for
XX improving gene targeting during further recombinant DNA protocols
XX involving plants. RAD2/FEN-1 endonucleolytic activity is essential
XX in DNA replication and nucleotide excision and repair reactions.
XX The exolytic activity is involved in double strand break repair and
XX end joining. The protein is also useful in strand exchange
XX reactions during homologous recombination. These functions may be
XX useful in gene targeting and in the production of male sterile
XX plants. The efficacy of gene targeting can be improved by the
XX overexpression of exogenous Rad2/FEN-1 while male sterile plants can
XX be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;
SQ
Query Match 97.8%; Score 1350.2; DB 21; Length 1463;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CGACCCGCGCTCGCGCCACAGCCGCGAGAGATGGGCATCAAGGTTTGAGCAAA 60
DB 49 CGCGCCCGCCACCGCCACAGCCGCGAGAGATGGGCATCAAGGTTTGAGCAAA 108
QY 61 CTGCTGGCGCAATGCGCCCAAGGCGATGAGAGAGCAAGATTGAGAGCTACTTCGGC 120
DB 109 CTGCTGGCGCAATGCGCCCAAGGCGATGAGAGAGCAAGATTGAGAGCTACTTCGGC 168
QY 121 CGCAAAATGCGCTGAGCGCAGATGAGCATATACCAATTCCTGATTTAGTTGGAAG 180
DB 169 CGCAAAATGCGCTGAGCGCAGATGAGCATATACCAATTCCTGATTTAGTTGGAAG 228
QY 181 ACAGGATGGAATCTCACAATGAGAGCTGAGTCACTAGTCAATTTGCAAGGATG 240
DB 229 ACAGGATGGAATCTCACAATGAGAGCTGAGTCACTAGTCAATTTGCAAGGATG 288
QY 241 TTCAACCGGACAAATTAAGTTTACTGGAAGCGGAGATCAAGCCAGTTTATGTTTGGATGCG 300
DB 289 TTCAACCGGACAAATTAAGTTTACTGGAAGCGGAGATCAAGCCAGTTTATGTTTGGATGCG 348
QY 301 AAGCTCTGATGATGAAGAAACAGAGCTTCTAAAGATACCTAAAGAGATGATGCA 360
DB 349 AAGCTCTGATGATGAAGAAACAGAGCTTCTAAAGATACCTAAAGAGATGATGCA 408

QY 361 ACCAAGATCTGACTGAGCAGTAGAGGTAGAGATAAAGATGCGATTTGAAAAATTTGAGC 420
DB 409 ACCAAGATCTGACTGAGCAGTAGAGGTAGAGATTAAGATGCGATTTGAAAAATTTGAGC 468
QY 421 AAGAGACTGTAAAGGTCTACACAGGCAACACACAGAAATTTTAACGGCTATTAGACTT 480
DB 469 AAGAGACTGTAAAGGTCTACACAGGCAACACACAGAAATTTTAACGGCTATTAGACTT 528
QY 481 ATGGGGGTCTCTGTTGTAGAGGCACTTCTGAAGCAGAAAGATGTGACGCCCTTTCG 540
DB 529 ATGGGGGTCTCTGTTGTAGAGGCACTTCTGAAGCAGAAAGATGTGACGCCCTTTCG 588
QY 541 ATAAACGATAAGTGTTCGCTGTTCCTCAGAAAGATATGAGCTCCCTTACTTTGGGCT 600
DB 589 ATAAACGATAAGTGTTCGCTGTTCCTCAGAAAGATATGAGCTCCCTTACTTTGGGCT 648
QY 601 CCAGGTTCTCTGCTCAATTAATGATCCAAAGTCCCAAGAAATACCTGTGATGAAAT 660
DB 649 CCAGGTTCTCTGCTCAATTAATGATCCAAAGTCCCAAGAAATACCTGTGATGAAAT 708
QY 661 GATGTGCCAAGGTTTGTGAGAGAGCTTGAACATCAACATGACCAAGTTCAATGATTTGTC 720
DB 709 GATGTGCCAAGGTTTGTGAGAGAGCTTGAACATCAACATGACCAAGTTCAATGATTTGTC 768
QY 721 ATCTGTGTGATGTGACTATTGTGATGATCAAGATGCGGGGGGCAACAGCTCTG 780
DB 769 ATCTGTGTGATGTGACTATTGTGATGATCAAGATGCGGGGGGCAACAGCTCTG 828
QY 781 AACTTATTCGTCACATGGGTCATGAAAGCATCTTGGAGATCTTAATAAGACAGA 840
DB 829 AACTTATTCGTCACATGGGTCATGAAAGCATCTTGGAGATCTTAAATAAGACAGA 888
QY 841 TATCAAAATTCCTGGAGAGCTGGCTTACCAAGAGCTGAGAGCTGTTTCAAGAGCCTAAT 900
DB 889 TATCAAAATTCCTGGAGAGCTGGCTTACCAAGAGCTGAGAGCTGTTTCAAGAGCCTAAT 948
QY 901 GTCACTTGTGATATTCCTGAGCTTAAATGAGCTGACCTGATGAGAGGGCTCATAGT 960
DB 949 GTCACTTGTGATATTCCTGAGCTTAAATGAGCTGACCTGATGAGAGGGCTCATAGT 1008
QY 961 TTCCGTGTAAGAAATATGTTTCAACGAAAGTGGGTGCAAGAGGCAATAGAGAGATC 1020
DB 1009 TTCCGTGTAAGAAATATGTTTCAACGAAAGTGGGTGCAAGAGGCAATAGAGAGATC 1068
QY 1021 AAATCTGCAAGAAATTAATGCTGCGCAAGGAGAGCTGAGTCTTTTCAAGCCACTGCG 1080
DB 1069 AAATCTGCAAGAAATTAATGCTGCGCAAGGAGAGCTGAGTCTTTTCAAGCCACTGCG 1128
QY 1081 ACCACATGACACCGCTTAAGGAGAGAGAGCTGCGATTAACCAAGCAAGGAGCTGCG 1140
DB 1129 ACCACATGACACCGCTTAAGGAGAGAGAGCTGCGATTAACCAAGCAAGGAGCTGCG 1188
QY 1141 AACAGAAAGAAAGAGGCTGTGGAAGAGAAATATCTTGATGCTTGTATGTAACATA 1200
DB 1189 AACAGAAAGAAAGAGGCTGTGGAAGAGAAATATCTTGATGCTTGTATGTAACATA 1248
QY 1201 CGACTAGCAAGAGCGGTGCGGATCCTGCTAGATTTATTAACCCCTGTGTTTA 1260
DB 1249 CGACTAGCAAGAGCGGTGCGGATCCTGCTAGATTTATTAACCCCTGTGTTTA 1308
QY 1261 ACTCAGAGCTTTGGTAAAGTTTCCCATGTTTCAAGCTGGGTAAAGTTAGTTGTGTTG 1320
DB 1309 ACTCAGAGCTTTGGTAAAGTTTCCCATGTTTCAAGCTGGGTAAAGTTAGTTGTGTTG 1368
QY 1321 AAGAGATTGTTGATCAAGATTAAGCAAACTTATGCGTGTGTTT 1363
DB 1369 AAGAGATTGTTGATCAAGATTAAGCAAACTTATGCGTGTGTTT 1411
RESULT 3
AAA27924
ID AAA27924 standard; cDNA; 1541 BP.
XX

[illegible]

Qy	181	ACAGCGATGGAAACTCTCCAAATGAAGCTGGTGAATGACTAGTCAATTTGGCAAGAAATG	240
Db	223	ACAGCGATGGAAACTCTCCAAATGAAGCTGGTGAAGTCACTACTGCTTTGGCAAGAAATG	282
Qy	241	TTTCAACCGGACATTAAGATTACTCTGAAGCGGGAAATCAAGCCAGTTATGTTTTGATGCG	300
Db	283	TTTCAACCGGACATTAAGATTACTCTGAAGCGGGAAATCAAGCCAGTTATGTTTTGATGCG	342
Qy	301	AAGCCTCTGTATGTAAGAACAAGAGCTTGGTAAAGATACTCAAAAAGAGATGATGA	360
Db	343	AAGCCTCTGTATGTAAGAACAAGAGCTTGGTAAAGATACTCAAAAAGAGATGATGA	402
Qy	361	ACCAAGATCTGACTGAGCGCATGTAAGGTAGAGATTAAGATGGGATTTGAAAATTTAGC	420
Db	403	ACCAAGATCTGACTGAGCGCATGTAAGGTAGAGATTAAGATGGGATTTGAAAATTTAGC	462
Qy	421	AAGGAGATGTAAAGTCACAGGCAACCAACAGAAATTTGAAACGGCTATTAGACTT	480
Db	463	AAGGAGATGTAAAGTCACAGGCAACCAACAGAAATTTGAAACGGCTATTAGACTT	522
Qy	481	ATGGGGTTCCTGTGTTGTAAGGCAACCTTCTGAAGCAGAAGATGTGACGCCCTTTGC	540
Db	523	ATGGGGTTCCTGTGTTGTAAGGCAACCTTCTGAAGCAGAAGATGTGACGCCCTTTGC	582
Qy	541	ATTAACGATTAAGGTGTCCGTGTTGGCTGCAGAAATATGATGATCCTTACTTTGGGGCT	600
Db	583	ATTAACGATTAAGGTGTCCGTGTTGGCTGCAGAAATATGATGATCCTTACTTTGGGGCT	642
Qy	601	CCACGGTTCCTGTGCATTTTAATGATCCAAGTTCCAGAAAATACCTGTATGGAATTT	660
Db	643	CCACGGTTCCTGTGCATTTTAATGATCCAAGTTCCAGAAAATACCTGTATGGAATTT	702
Qy	661	GATGTGGCAGATTTTGGAGAGAGCTGAACATCCATGAGACAGTCAATGATTTGGC	720
Db	703	GATGTGGCAGATTTTGGAGAGAGCTGAACATCCATGAGACAGTCAATGATTTGGC	762
Qy	721	ATCCGTGTGATGTGACTATTTGATATGATCAAGATATCGGGGGGCAACAGCTCTG	780
Db	763	ATCCGTGTGATGTGACTATTTGATATGATCAAGATATCGGGGGGCAACAGCTCTG	822
Qy	781	AAACTTATTTCTCAACATGGGTCCTATGAAAGATCTTGAGAAATCTTAATAAGACAGA	840
Db	823	AAACTTATTTCTCAACATGGGTCCTATGAAAGATCTTGAGAAATCTTAATAAGACAGA	882
Qy	841	TATCAATTTCTGAGAGCTGCGCTTACCAGAAAGCTCGAGCGTTGTTCAGAGGACTTAAT	900
Db	883	TATCAATTTCTGAGAGCTGCGCTTACCAGAAAGCTCGAGCGTTGTTCAGAGGACTTAAT	942
Qy	901	GTGCATTGATATTTCTGAGCTTAAATGATGATGCACTGATGAGAGGCTCTTAATGT	960
Db	943	GTGCATTGATATTTCTGAGCTTAAATGATGATGCACTGATGAGAGGCTCTTAATGT	1002
Qy	961	TTTCCGTGTAAGATTAATGTGTTCAACGAAGATCGGGTGCAGAAAGGCCATAGAGAAATC	1020
Db	1003	TTTCCGTGTAAGATTAATGTGTTCAACGAAGATCGGGTGCAGAAAGGCCATAGAGAAATC	1062
Qy	1021	AAATCTGCCAAGATTAATCTGTCGCAAGAGAGCTGAGTCTTTTTCAGAGCAACTGCC	1080
Db	1063	AAATCTGCCAAGATTAATCTGTCGCAAGAGAGCTGAGTCTTTTTCAGAGCAACTGCC	1122
Qy	1081	ACCAATCATGCAACCGCTTAAACGGAAAGAGCTTCGATTAACCAAGCAAGGCACTCTCG	1140
Db	1123	ACCAATCATGCAACCGCTTAAACGGAAAGAGCTTCGATTAACCAAGCAAGGCACTCTCG	1182
Qy	1141	AACCAAAAAACAAGGCTGTGTGGAAGAAAGAAATTAATCTTGATGCTTGATGTACACTA	1200
Db	1183	AACCAAAAAACAAGGCTGTGTGGAAGAAAGAAATTAATCTTGATGCTTGATGTACACTA	1242
Qy	1201	CGACTACGAAGACAGCGGTGGGTGATACACTTCGCTTAGATTTTAACTACCCCTGTTTTA	1260
Db	1243	CGACTACGAAGACAGCGGTGGGTGATACACTTCGCTTAGATTTTAACTACCCCTGTTTTA	1302

QY 1261 ACTCAGAGCTTTGTTAAAGTTCCGCATGTTTCAAGCTGGGGTAAGTACTGTTGTTG 1320
DB 1303 ACTCAGAGCTTTGTTAAAGTTGCTCATGTTTCAAGCTGGGGTAAGTACTGTTGTTG 1362
QY 1321 AAGAGATTGTTTACCAAGTAACAACCTTATCGCTGTTT 1363
DB 1363 AAGAGATTGTTTACCAAGTAACAACCTTATCGCTGTTT 1405

RESULT 4
AA27926

ID AAA27926 standard; cDNA, 1478 BP.

XX AAA27926;

DT 12-SEP-2000 (first entry)

XX Maize Rad2/FEN-1 cDNA.

KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
endonuclease; exonuclease; DNA repair; gene targeting; ss.

XX Zea mays.

FH Key Location/Qualifiers

FT CDS 97..1236

XX WO200036109-A1.

XX 22-JUN-2000.

XX 16-NOV-1999; 99WO-US27147.

XX 15-DEC-1998; 98US-0112332.

XX (PION-) PIONEER HI-BRED INT INC.

XX Mahajan PB;

XX WPI; 2000-452026/39.

XX P-PSDB; AAY95310.

PT Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
recombination and repair in transgenic plants, e.g. for gene targeting
and the production of male sterile plants.

XX Example 1; Page 77-79; 85pp; English.

XX The present sequence is that of maize cDNA coding for RAD2/FEN-1
(see AAY95310). The cDNA was isolated from a library prepared
from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a
structure specific endonuclease which under certain conditions also
acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
The protein is involved in the regulation of DNA repair and
recombination in plant systems and therefore may be used for
improving gene targeting during further recombinant DNA protocols
involving plants. Rad2/FEN-1 endonucleolytic activity is essential
in DNA replication and nucleotide excision and repair reactions.
The exolytic activity is involved in double strand break repair and
end joining. The protein is also useful in strand exchange
reactions during homologous recombination. These functions may be
useful in gene targeting and in the production of male sterile
plants. The efficacy of gene targeting can be improved by the
overexpression of exogenous Rad2/FEN-1 while male sterile plants can
be produced by the down regulation of Rad2/FEN-1 expression.

XX Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;

Query Match 96.6%; Score 1334.2; DB 21; Length 1478;

Best Local Similarity 98.7%; Pred. No. 0;
Matches 1345; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 CGACCCAGCGCTCCGGCCACAGCCGCCGACGAGATGGCATCAAGGTTTGACGAA 60
DB 61 CGCGCCCGCCACCCGCCACAGCGCCGCCGACGAGATGGCATCAAGGTTTGACGAA 120
QY 61 CTGCTGGCGGCAATGCGCCCAAGGCGATGAAGAGACAGAGTTCGAGAGCTTCGCG 120
DB 121 CTGCTGGCGGCAATGCGCCCAAGGCGATGAAGAGACAGAGTTCGAGAGCTTCGCG 180
QY 121 CGCAAAATCGCCGTCGACGCCAGCATATACCATGCTCTGATTTGAGAGG 180
DB 181 CGCAAAATCGCCGTCGACGCCAGCATATACCATGCTCTGATTTGAGAGG 240
QY 181 ACAGGATGAAATCTCAGCAAAATGAAGCTGTGAAGTCACTAGTCATTGCAAGAAATG 240
DB 241 ACAGGATGAAATCTCAGCAAAATGAAGCTGTGAAGTCACTAGTCATTGCAAGAAATG 300
QY 241 TTCAACCGGCAATATACATTTACGGAAGCGGGAATCAAGCCAGTTATGTTTATGATG 300
DB 301 TTCAACCGGCAATATACATTTACGGAAGCGGGAATCAAGCCAGTTATGTTTATGATG 360
QY 301 AAGCTCCTGATATGAAGAAACAAGAGCTTGCTAAAGATCTCAAAAAGAGATGATGCA 360
DB 361 AAGCTCCTGATATGAAGAAACAAGAGCTTGCTAAAGATCTCAAAAAGAGATGATGCA 420
QY 361 ACCAAAGATCTGACGAGCGAGTGAAGCTTGAAGATTAAGATCGCTTTAAAAATTTAGC 420
DB 421 ACCAAAGATCTGACGAGCGAGTGAAGCTTGAAGATTAAGATCGCTTTAAAAATTTAGC 480
QY 421 AAGAGACTGTAAGAGTGCACAGGCAACACACAGAGATTTGAACGGCTATTGAAGACTT 480
DB 481 AAGAGACTGTAAGAGTGCACAGGCAACACACAGAGATTTGAACGGCTATTGAAGACTT 540
QY 481 ATGGGGTTCCTGTTGTAAGAGCACTTCTGAAGCAAGCAAGTACAGCTTCGCTTGC 540
DB 541 ATGGGGTTCCTGTTGTAAGAGCACTTCTGAAGCAAGCAAGTACAGCTTCGCTTGC 600
QY 541 ATAAAGATTAAGGTTTGGCTTCTCAAGAAATATGAGTCCCTTACTTTGGGGCT 600
DB 601 ATAAAGATTAAGGTTTGGCTTCTCAAGAAATATGAGTCCCTTACTTTGGGGCT 660
QY 601 CCAGGTTCTCTGCTCATTTATGATGATCCAAAGTTCAGAAATATGATGATGATG 660
DB 661 CCAGGTTCTCTGCTCATTTATGATGATCCAAAGTTCAGAAATATGATGATGATG 720
QY 661 GATGTTGCCAAGGTTTGGAGAGCTTGAATCAATGAGCAGTTCAATGATTTGTC 720
DB 721 GATGTTGCCAAGGTTTGGAGAGCTTGAATCAATGAGCAGTTCAATGATTTGTC 780
QY 721 ATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 781 ATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 781 AAAGTATTTGCTCAACATGGTTCATGAAGAAAGCATCTTGGAGATCTTAATAAAGCACA 840
DB 841 AAAGTATTTGCTCAACATGGTTCATGAAGAAAGCATCTTGGAGATCTTAATAAAGCACA 900
QY 841 TTTCAAATTCCTGAGAGTTCCTTACCAAGAAAGCTGAGGCTTGTCAAGAGCCTAAT 900
DB 901 TTTCAAATTCCTGAGAGTTCCTTACCAAGAAAGCTGAGGCTTGTCAAGAGCCTAAT 960
QY 901 GTTCATTTGATATCTCTGAGCTTAATATGACCTGACCTATGAGAGGCTCATTAAGT 960
DB 961 GTTCATTTGATATCTCTGAGCTTAATATGACCTGACCTATGAGAGGCTCATTAAGT 1020
QY 961 TTCTGTGTAAGAAATATGATTTCAAGAGATGGGTGCAAGGCAATAGAGATG 1020
DB 1021 TTCTGTGTAAGAAATATGATTTCAAGAGATGGGTGCAAGGCAATAGAGATG 1080
QY 1021 AAATCTGCCAAGATTAATCTGTCAGAGAAAGCTGAGTCTTTTCAAGCAACTGCC 1080
DB 1081 AAATCTGCCAAGATTAATCTGTCAGAGAAAGCTGAGTCTTTTCAAGCAACTGCC 1140

QY 1081 ACCACATCAGCAGCCGCTAAACGGAGAGACTTCGGATAAACAAAGCAGCTGCG 1140
1141 ACCACATCAGCAGCCGCTAAACGGAGAGACTTCGGATAAACAAAGCAGCTGCG 1200
QY 1141 AACAGAGAAACAAAGGCTGTGGAGAGAGAAATATCTTGATGCTGTACACACTA 1200
1201 AACAGAGAAACAAAGGCTGTGGAGAGAGAAATATCTTGATGCTGTACACACTA 1260
QY 1201 CGACTCAGAAAGCAGCGGTGGCGTATCCTTCCTAGATTATTAACCTCCCTGTTTA 1260
1261 CGACTCAGAAAGCAGCGGTGGCGTATCCTTCCTAGATTATTAACCTCCCTGTTTA 1320
QY 1261 ACTCAGAGCTTTGGTAAAGTTCCGCCATGTTTCAAGCTGGGGTAGTTAGTTGGTTTG 1320
1321 ACTCAGAGCTTTGGTAAAGTTCCGCCATGTTTCAAGCTGGGGTAGTTAGTTGGTTTG 1380
QY 1321 AAGAGATTGTGTACCAAGTAAACAAACTATCCCTGTTT 1363
1381 AAGAGATTGTGTACCAAGTAAACAAACTATCCCTGTTT 1423
Db

RESULT 5
AA02117
ID AA02117 standard; DNA; 1144 BP.

XX AAX02117;

XX 23-APR-1999 (first entry)

DE Human FEN-1 DNA fragment.

KM FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
neoplasia; antineoplastic agent; cleavage; ss.

OS Homo sapiens.

PN US5874283-A.

PD 23-FEB-1999.

PF 30-MAY-1995; 95US-0455968.

PR 30-MAY-1995; 95US-0455968.

PA (HARR/) HARRINGTON J J.

PA (HSIEH/) HSIEH C.

PI (LIEB/) LIEBER M R.

PI Harrington JJ, Hsieh C, Lieber MR;

DR WPI; 1999-179985/15.

PT DNA encoding flap endonuclease polypeptides - useful for producing

PS e.g. recombinant polypeptides

PS Claim 3; Column 25; 58pp; English.

XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC a novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.

XX Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other;

QY Query Match 21.9%; Score 302.6; DB 20; Length 1144;

Best Local Similarity 56.6%; Pred. No. 1.5e-77; Mismatches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 37 ATGGGATCATAGGTTTGTGACGAACCTGCTGCGGACATATGCGCCCAAGGATGAAGAG 96
1 ATGGGAATTCGAAGGCTGCGCAAACTAATGCTGATGTTGGCCCGCCAGTGCATTCGGAG 60
Db
QY 97 CAGAACTTGAGAGACTACTTGGCGGCAAAATCGCGCTGCGACGACACATGACATATAC 156
61 AATGACATCAAGAGCTACTTGGCGGCAAAATCGCGCTGCGACGACATGACATATAT 120
Db
QY 157 CAGTTCCTGATTTGATTTGAGAGAGACAGCATGGAACCTGCTCACAATGAGGCTGTGA 216
121 CAGTTCCTGATTTGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 177
QY 217 GTCACTAGTCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276
178 ACCACAGGCGACCTGATGGGATGTTTACCGACACATGCTCAGATGATGAGAAAGGATC 237
QY 277 AAGCAGATTATTTGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 336
238 AAGCAGATTATTTGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297
QY 337 AGATCTCAAAAAGAGATGATGCAACCAAGATGCTGACTGAGGAGATGATGAGAT 396
298 CGCAGTACGCGCGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 357
QY 397 AAGATGCGATTGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
358 GAGCAGAGAGTGAATAAATTTCACTAAGCGGCTGTGAGAGTCACTAAGCAGCAATGAT 417
QY 457 GATTGTAACGGCTATTAGACTTATGAGGCTCTGTTGTAAGAGAGAGAGAGAGAGAG 516
418 GAGTGCATAATCTGCTGAG 477
QY 517 GAAGCAGATGTCAG 576
478 GAGGAGAGAGTGTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537
QY 577 ATGAGCTCCCTTACTTTTGGGCTCCAGGCTTCTTCTGCTATTATGATCCAGTTC 636
538 ATGAGCTCCCTTACTTTTGGGCTCCAGGCTTCTTCTGCTATTATGATCCAGTTCAG 597
QY 637 AAGAAATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
598 AAAAAGCTGCAATCCAGGAATTTCCAGTCCAGGCTGATGATGATGATGATGATGATGAT 657
QY 697 ATGACCAATTCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 756
698 CAGGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
QY 757 GGTATCGGGGGGCAACAGCTGGAACCTTATTTGATGATGATGATGATGATGATGATGAT 816
718 GGTATTTGGGCGGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
QY 817 TTGAGATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 876
778 GTGCGGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
QY 877 CGAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 933
838 CACCACTCTTCTTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
QY 934 GCACCTGATGAGAGAGGCTCATTAAGTTCCTGTAAGAGATTAAGTTCACAGAGAT 993
898 GAGCAATTAAGAGAGAGGCTCATTAAGTTCCTGTAAGAGATTAAGTTCACAGAGAT 957
QY 994 CGGGTGAAGAGAGGCTCATTAAGTTCCTGTAAGAGATTAAGTTCACAGAGAGAT 1053
958 CGAATTCGAGAGAGGCTCATTAAGTTCCTGTAAGAGATTAAGTTCACAGAGAGAT 1017
QY 1054 CTCGATCTTTTCAAGCAAGCTGCGACACATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1094
1018 CTGATGATTTTCAAG 1058


```

RESULT 6
ID AAX02107 standard: cDNA: 1144 BP.
XX
AC AAX02107;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 cDNA.
XX
KM FEN-1: human; flap endonuclease; detection: diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1144
FT /tag="a
FT /product="FEN-1"
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
DR WPI: 1999-179985/15.
DR P-PSDB; AAM92504.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
PS e.g. recombinant polypeptides
XX
PS Disclosure: Fig 1B; 58pp: English.
XX
CC This sequence encodes a human FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other;
XX
Query Match 21.9%; Score 302.6; DB 20; Length 1144;
Best Local Similarity 56.6%; Pred. No. 1.5e-77;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

```

```

Db 178 ACCACAGCCACCTGATGGGATGTTCTACCGCACCATTTGTCATGATGAGAAAGCGATC 237
QY 277 AAGCCAGTTATGTTTGTATGGCAAGCCTCCTGATATGAAAGAAACAAGCCTTGCTAA 336
Db 238 AAGCCGCTGATGCTCTTGTATGGCAAGCCTCCTGATGATGAGCAGCTTGCGCCAAA 297
QY 337 AGATACTCAAAAAGAGATGTCACCAACCAAGATCTGACTGAGGAGTAGAGTAGAGAT 396
Db 298 CCGAGTGAAGCGCGCGGCTGAGCAGAGAAGCAGCTGACGAGGCTCAGGCTGCGGGGCC 357
QY 397 AAGATGCCATTGAATAATTTGAGCAAGAGACTGTAAAGTCAACAGGCAACAGCAAG 456
Db 358 GAGCAGAGAGGTGGAAATTCACTAAGCGGCTGTGAAGCTACTAAGCAGCAATGAT 417
QY 457 GATTGTAAAGGCGCTATTAAGACTATAGGAGGCTCTCTGTGTAGAGGCACTTCTGAAGCA 516
Db 418 GAGTGCACCAATCTGTGTAGCGCTCATGGGCATCCCTTATCTTGATGCAACCCAGTAGCA 477
QY 517 GAAGCAGAAATGTCAGCCCTTTGCATTAACGATAGGTGTTGCTGTTCAGAAAGAT 576
Db 478 GAGGCCAGCTGTGCTGCCCTGGTGAAGGCTGCAAAAGCTATGCTGGGCTACCGAGGAC 537
QY 577 ATGAGCTCCCTTACTTTTGGGCGTCCACGCTTCTCTGCTCATTTAATGATCCAAAGTTC 636
Db 538 ATGAGCTGCTCACCCTTGCGGAGCCCTGTGTAATGACACACGACGACGCGCAGTAGAGCC 597
QY 637 AAGAAATATCCTGTGATGTAATTTGATGTTGCCAAGGTTTGGAGAGCTTGAAGCTACC 696
Db 598 AAAAAGCTCCCAATCAGGAATTCACCTGAGAACCCGATTCGCAAGGCTGGGCGTGAAC 657
QY 697 ATGAGCAGCTTATGATTTGATTTGTCATTCCTGATGATGATGATGATGATGATGATGAT 756
Db 658 CAGGAAACATTTGTGATGTGATGCTGATTCCTGATGATGATGATGATGATGATGATGAT 717
QY 757 GGTATCGGGGGGCAACAGACTGTGAACCTTATTCGTCAACATGGGTCCATAGAAAGATC 816
Db 718 GGTATTTGGGCCCAAGGGGCTGTGACCTCATCCAGAAAGCAAGAGACTGAGAGATC 777
QY 817 TTGAGAAATCTTAATTAACAGATATCAAAATTCCTGAGGACTGGCTTACCAAGAGCT 876
Db 778 GTGCGGCGACTTGACCCCAACAAGTACCTGTGCGCAAGAAATTTGGCTCCCAAGAGAGCT 837
QY 877 CGACGCTTTGTAAGAGGCTTAATG---TCACATTTGATATTTCTGACCTTAATGAGCT 933
Db 838 CACCACTCTTTCTTGAACCTGAGGTGCTGACCCAGAGCTGTGAGAGCTGAAGTGAAGC 897
QY 934 GCACCTGATGAGAGGCTGCTCATAGTTTCCGTGTAAGAAAGATATGTTCAACAGAGAT 993
Db 898 GAGCCAAATGAGAGAGCTGATCAAGTTCAAGTGTGATGTAAGAAAGAGCTTCTGAGAG 957
QY 994 CGGGTACAAAGGCTATGAGAGATCAAAATCTGCCAAGAAATTAATGCTGCCAAGAGAGA 1053
Db 958 CGAATCCGACGTGGGCTCAAGAGAGCTGAGTAAGAGCCGCCAAGGAGAGCACCAGGGCGC 1017
QY 1054 CTCGAGTCCCTTTTCAAGCACTGCCACACACATGAGCAGC 1094
Db 1018 CTCGATGATTTCTTCAAGGTGACCGGCTCAGCTCTTCAAGC 1058

```

```

RESULT 7
ID AAX02111 standard: DNA: 2033 BP.
XX
AC AAX02111;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 genomic DNA.
XX
KM FEN-1: human; flap endonuclease; detection: diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; ss.
XX

```

```

OS Homo sapiens.
XX key Location/Qualifiers
FH CDS 104..1240
FT /*tag= a
FT /product= "FEN-1"
XX
XX US5874283-A.
XX
XX 23-FEB-1999.
XX
XX 30-MAY-1995; 95US-0455968.
XX
XX 30-MAY-1995; 95US-0455968.
XX
XX (HARR/) HARRINGTON J J.
XX (HSIEH/) HSIEH C.
XX (LIEBER/) LIEBER M R.
XX
XX Harrington JJ, Hsieh C, Lieber MR;
XX
XX WPI: 1999-179985/15.
XX P-PSDB: AAM92508.
XX
XX DNA encoding flap endonuclease polypeptides - useful for producing
XX e.g. recombinant polypeptides
XX
XX Disclosure: Fig 5A-B; 58pp; English.
XX
XX This sequence encodes a human FEN-1 (flap endonuclease) protein. This
XX protein can be used in methods for detecting a pathological condition in
XX a patient, for diagnostic purposes, for screening for antineoplastic
XX agents and carcinogens, for diagnostic staging of neoplasia, for
XX producing recombinant flap endonuclease for use as research or
XX diagnostic reagents, for producing antibodies reactive with the novel
XX polypeptides, for producing transgenic nonhuman animals expressing the
XX novel polypeptides encoded by a transgene. The invention also provides
XX novel molecular cloning techniques and reagents involving cleavage of
XX a flap or nick with a flap endonuclease.
XX
XX Sequence 2033 BP; 500 A; 493 C; 579 G; 461 T; 0 other;
SQ
XX
XX Query Match 20.0%; Score 276.6; DB 20; Length 2033;
XX Best Local Similarity 55.9%; Pred. No. 7.4e-70;
XX Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;
XX
XX 37 ATGGGCAATCAAGGCTTGGAGAACTGCTGGCGGACATGCGCCCAAGCGGATGAAGAG 96
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 104 ATGGAATTCACGGCCCTTGCCAACTAATGCTGATGTGTGCCCCCGCCAGTCCGCTGAG 163
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 97 CAGAACTTCAGAGACTCTTCGGCGGCAAAATGCGCGTGCAGCGCAGCATGAGCATATAC 156
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 164 AATGACATCAAGAGACTTGTGTCGCAAGTGCGCATTCGATCCCTCCATGAGCATCTAC 223
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 157 CAGTTCCTGATTTGATTTGGAGAGACAGGCACTGAAACTCTCAAAATGAAGCTGTGAA 216
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 224 CAGTTCCTGATTTGCTGTCTCAG--GTTGGGATGTCTGCAAGAACGAGGGGTGAG 280
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 217 GTCACATGCTATTTCAGAGAGATGTTCAACCGGCAATTAAGATTATTCGAGAGGGGAATC 276
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 281 ACCACGAGCC--TGATGGGCAATGTTCACTGATACCATGCG--CATGGAATGGGATTC 334
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 277 AAGCCAGTTATGTTTGGATGGCAGCTCTGATATGAAGAAACAAAGAGCTTGTAA 336
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 335 AAGCCTGTGTGCTTTGATGGCAACACACAGCTGAAGTCAAGCGAGGCTGGCCAG 394
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 337 AGATACCTCAAAAAGATGATGATCAACCAAGATCTGACTGAGGCACTGAAGGTAGAGAT 396
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 395 CGCAGTGAGAGGCGCGCGAGGCTGAGAAAGCACTGCAGAGGCTCAGAGCTGGGATG 454
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 397 AAAGATGCGATTGAAAAATTGACAGAGAGGACTTAAGGTCAACAAAGCAACCAAGAA 456
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 455 GAGGAGGAGGTGAGAAAGTTCAACCAAGAGGCTGTGAAGGTCCACCAAGCAACCAATGAT 514

```

```

XX 457 GATTGAACGGCTATTAGACTTATGGGGTTCCCTGTTGTGAGGACACCTTCAGACA 516
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 515 GAGTGCAAACACCTGCTGAGCCTCATGCTCCCTTACCTGATGACACCGCAGCA 574
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 517 GAGCAGAAATGTGCGACCGCTTTGACATAAACGATAAGGTGCTGTTGCTTCAGAAAT 576
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 575 GAGGCCAGCTGTGCTCCTCGGCAAAAGGTGCAAAAGTATGCTGCGGCCACGAGGAC 634
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 577 ATGACCTCCCTTACTTTTGGGGCTCCACGCTTCCTTCGTCATTATATGATCCAGTTC 636
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 635 ATGACCTGCTCCTACTTTTGGCAGCCCGCTGCTAATGCGACCTTAACCTCAGTAGGCC 694
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 637 AAGAAATACCTGTGATGATGATTTGTCAGAGCTTGTGAGAGAGCTTAAGTCACTCAC 696
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 695 AAGAAAGCTGCCATCCCAAGATTCATCTGAGCGGCTCTCGAGAGAGTGGGTGGAAC 754
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 697 ATGACCACTTCATTTGATTTGTGATCTCGTGTGATGATGATGATTTGATGATCAAA 756
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 755 CAGGAGCAGTTTGTGATCTGTGCATCTCTGCTGATGAGCACTGCGAGAGATCCGT 814
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 757 GGTATCGGGGGGCAACAGCTCTGAACCTTATTCGTCACATGCGTCCATPAGAAGCATC 816
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 815 GCGATTGCGCCCAAGCGGCTGTGATCTCATCAGAAACATAGAGCATCGAGAGATC 874
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 817 TTGGAGATCTTATATAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGAGCT 876
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 875 GTGAGCGGGCTGAGACCCGACAGTACCTCCGTTCCAGAGAACTGGCTCCACAGAGAAC 934
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 877 CGAGCGCTTGTCAAGAGAGCTTAA--TGTCACATTTGATATTCCTGAGCTAAATGAGCT 933
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 935 CAGCAGCTCTTCCCTGAGAGCCAGAAAGTGTGAGCCAGCAAGTCTGTGAGTGAAGTGAGC 994
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 934 GCACCTATGAGGAGGGCTCATAGTTCTCTGTAAGATATATGATGTTTCAAGAGAT 993
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 995 GAGCCAAATGAGAAAGAGTGTGTAATTTATGTGTGTAAGAAAGCTTTTGAAGAG 1054
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 994 CGGCTGACAAAGGCCATAGAGAGATCAAAATCTGCCAAGATTAATGTGCCAAGAGAG 1053
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 1055 CGAATTCGACGTGGGCTCAACGCGCTGATGAAGAGCCGCCAGGCGACACCCAGGGAGC 1114
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 1054 CTCGAGTCTTTTCAAGCCACAGTGCACCAACATCAGACCGCTTAAACGGAAGAGACT 1113
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 1115 CTCGATGATTTCTTCAAGGTGACAGGCTCACTCTCTCAGC-----TAAACGC 1162
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 1114 TCGATTAACAGCAGAGGAGCTGCGAACAAGAAACAAAGGCTGTGTAAGAGAGAAA 1173
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 1163 AAGGAGCCAGAAACCAAGAGGCGCTGCTGTAAGAAAGCAAGACTGGGGGAGCGGGAAG 1222
XX ||||| || || || || || || || || || || || || || || || || || || ||
XX 1174 T 1174
XX 1223 T 1223
XX
XX RESULT 8
XX ABL14287
XX ID ABL14287 standard; cDNA; 1300 BP.
XX
XX ABL14287:
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 37343.
XX
XX Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX

```

PE 23-MAR-2001: 2001WO-US09231.
 XX 23-MAR-2000: 2000US-191637P.
 PR 11-JUL-2000: 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li FMD, Myers EW;
 XX
 DR MPI: 2001-656860/75.
 DR P-PSDB: ABB70184.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS
 CC Claim 1; SEQ ID NO 37343; 21bp + Sequence Listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB101840-AB116175).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 CC
 CC Sequence 1300 BP; 346 A; 313 C; 369 G; 272 T; 0 other;
 SQ
 Query Match 20.0%; Score 276; DB 23; Length 1300;
 Best Local Similarity 55.2%; Pred. No. 8.7e-70;
 Matches 581; Conservative 0; Mismatches 465; Indels 6; Gaps 2;
 QY 37 ATGGGCATCAAGGTTGACGAACCTGCTGGCGACATGCGCCCAAGCGCATGAAGAG 96
 DB 1 ATGGGAATTTTGGCTATCTAAGCTCATTTGCGCGTGGCCGACAGGATTCGCCAA 60
 QY 97 CAGAAGTTCGAGAGCTACTTCCGCCGCAAAATCGCCGTCAGCGCCAGCATATATAC 156
 DB 61 AGTAGATGAAGCATTTTTCGTCGCAAGTAGCAATGATGATGCTGTATGCTGTATC 120
 QY 157 CAGTTCCTGATTTGATTTGGAAGACAGCATGAAATCTCACAATATACCTGTGA 216
 DB 121 CAGTTCCTGATTCGCGGCGCTCCGAAGGC--GCCAGTTGGCCACCGTAATGTGAT 177
 QY 217 GTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACCAATTAAGATTACTGAAACGGGAATC 276
 DB 178 CCCACGTCGCCACTAATGCGCATGTTCTACGCGCACCATCGATGCTGACAAAGGAATC 237
 QY 277 AAGCGAGTTTATGTTTGAAGCAAGCCCTGATGATGAAGCAAGAGTTGCTTAA 336
 DB 238 AAGCCGGATATGTCCTTCGACGGAAGCCACAGATCTCAAGTCCGCTGAGCTGGCAG 297
 QY 337 AGATTACTCAAAAGAGATGATGCAACCAAGATGATGAGGAGCTAGAGGTAGAGAT 396
 DB 298 CGCGCGCGCGCGCGGAGGAGGAGCACTGAAGCGCGCCACCGCATCGGAGAT 357
 QY 397 AAGATCGATTTGAAATATGAGCAAGAGAGCTAAAGGTCAAGGCAAGCAACGA 456
 DB 358 GATGCGGATTCGAAAGTTTAATCGCCGATTTGTCGCGGTAAAGAGAGAGACCCCAA 417
 QY 457 GATTGTAAAGCGCTATTAAACATTATGAGGCTCTGTTGTAGAGGACCTTCTAGACA 516
 DB 418 GAGGCCAAGGAAGTGTACACTTAATGAGTGGTGCCTTAATGATGACCGGCAAGCG 477
 QY 517 GAAGCAGATGTGACGCCCTTTGATTAAGATTAAGATGCTGCTGTTCTCAGAAAT 576
 DB 478 GAGGCCAGGTGTGCGCTGTGTAAGGTGAAGGTTATGCCACCGCAGGAGAT 537
 QY 577 ATGAGCTCCCTACTTTTGGGGCTCCACGCTGCTCTTGTGATTAATGATCAAGTTCC 636

DB 538 ATGATGCCCTTACATTTGGAATCTACAAACTGTTGATACCTTACCTTACAGCGGCA 597
 QY 637 AAGAAATACCTGTGATGAAATTTGATGTCGAAGTTTGGAGAGCTTGAATCACC 696
 DB 598 CGAAGATCCCTCAAGAGTTCAGTACGACAAAGCTGTGGAAGGTGCGCCATTAA 657
 QY 697 ATGACACAGTTCATTTGATTTGATCTGCTGTGATGATGATGATGATGATGATGAT 756
 DB 658 AATCGAGCTTATTTGATCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
 QY 757 GGTATCGGGGCGCAACAGCTCTGAACCTTAATGCTCAACATGCTTACATGAAGCAT 816
 DB 718 GGTATGACCCCAACAGCGATGATGATGATGATGATGATGATGATGATGATGAT 777
 QY 817 TTGAGAACTCTAATTAAGACATATCAATTCCTGAGCTGCTGCTGCTGCTGCTGCT 876
 DB 778 CTGATTAACCTTGACTTACCAATATACACCTGCGCCGAGAACTGAACTCAAGGTGCG 837
 QY 877 CGACGCTTTCACAGACCTTAATGTCAC--ATTGATATTCCTGAGCTAAATGACT 933
 DB 838 CGGGAATCTTTCATGCAACCGAGGTACTGATGCTGCTGCTGCTGCTGCTGCTGCT 897
 QY 934 GCACCTGATGAGAGGCTCATATAGTTTCTGTTAAAGATTAATGTTTCAACGAAGAT 993
 DB 898 GACCGGATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
 QY 994 CGGTCGCAAGGCGCATGAGAGATCAATTCCTGCAAGATTAATGCTGCAAGAGAG 1053
 DB 958 CGGTCGCAAGGCGCATGAGAGATCAATTCCTGCAAGATTAATGCTGCAAGAGAG 1017
 QY 1054 CTCGAGTCTTTTTCAGGCAACTGCCACAC 1085
 DB 1018 CTCGATAGCTTCTTTAAGACACTGCCGAC 1049
 RESULT 9
 AA02108
 ID AA02108 standard; cDNA; 1930 BP.
 XX
 AC AA02108;
 XX
 DT 23-APR-1999 (first entry)
 XX
 DE Mouse FEN-1 cDNA.
 XX
 KW FEN-1; mouse; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasma; antineoplastic agent; cleavage; ss.
 OS Mus sp.
 PN US874283-A.
 PD 23-FEB-1999.
 XX
 PF 30-MAY-1995; 95US-0455968.
 XX
 PR 30-MAY-1995; 95US-0455968.
 XX
 PA (HARR/) HARRINGTON J J.
 PA (HSTE/) HSTEH C.
 PA (LIEB/) LIEBER M R.
 PI Harrington J, Hsieh C, Lieber MR;
 DR MPI: 1999-179985/15.
 DR P-PSDB: AAW92505.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 PS Disclosure; Fig 2B; 58pp; English.
 XX

CC This sequence encodes a mouse FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing antibodies reactive with the novel
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.

XX Sequence 1930 BP; 482 A; 459 C; 553 G; 436 T; 0 other;

Query Match 19.7% Score 272.2; DB 20; Length 1930;

Best Local Similarity 55.1%; Pred. No. 1.4e-68; Mismatches 488; Indels 24; Gaps 4;

Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;

QY 37 ATGGGCAATCAGGCTTGGAGCAACTCTGCGGACAAATGCGCCGAGAGCGGATGAAGAG 96
 DB 1 ATGGGCAATCAGGCTTGGAGCAACTCTGCGGACAAATGCGCCGAGAGCGGATGAAGAG 60
 QY 97 CAGAGTTCAGAGCTACTTCCGCGCAAAATCGCCGTCGAGCGGACGATGAGCATATAC 156
 DB 61 AATGACATCAGAGCTACTTGGCTAAAGTGGCCATCGATGCTCATGAGCATCTAC 120
 QY 157 CAGTTCCTGATTAGTGGAGACAGGATGAACTCTCACAAATGAAGCTGTGA 216
 DB 121 CAGTTCCTGATTAGTGGAGACAGGATGAACTCTCACAAATGAAGCTGTGA 177
 QY 217 GTCCTGATTAGTGGAGACAGGATGAACTCTCACAAATGAAGCTGTGA 276
 DB 178 ACCGACGAGCTGATGGGATGTTAT-----GGCAACCATCCGATGAGAAATGGCATC 231
 QY 277 AAGCCATTATGTTTGTATGCAAGCTCCTGTATGAAGAAACAAGCTTGTAA 336
 DB 232 AAGCTGTATGCTTTGTATGCAAGCTCCTGTATGAAGAAACAAGCTTGTAA 291
 QY 337 AGATCTCAAAAGATGATGCAACCAAGATGCTAGTGGAGAGTGAAGTGAAGAT 396
 DB 292 CGCAGTGAAGGCGCGCGGAGGCTGAGAGCAACTGCAAGGCTGAGAGGCTGGATG 351
 QY 397 AAGATCGCTTAAATAATTTGAGCAAGAGCTTAAAGGTCAAGCAACCAAGAA 456
 DB 352 GAGGAGAGGTGAGAGATTCACCAAGAGGCTGTAAGGTCAACCAAGAAATGAT 411
 QY 457 GATTCTAAAGGCTATTAGACTTATGGGCTTCTGTTGAGAGCACTTCTGAAGA 516
 DB 412 GAGTCAAAACACCTCGTAGGCTCATGGGCTTACCTTGAATGACCCAGAGGACA 471
 QY 517 GAAGCAAAATGTCAGAGCTTTCATAAAGATTAAGGTTCGCTTCAAGAT 576
 DB 472 GAGGCAAGCTGTGCTGCGGCAAAAGGCTGCAAAAGTCTATGCTGCGGCGAGGAG 531
 QY 577 ATGAGCTCCCTTCTTGGGCTCCACGCTTCTTCATTTAATGATCAAGTTC 636
 DB 532 ATGAGCTCCCTTCTTGGGCTCCACGCTTCTTCATTTAATGATCAAGTTC 591
 QY 637 AAGAAATACCTGTGATGAAATTTGTCGCAAGGTTTGGAGAGCTGAAGTCAAC 696
 DB 592 AAGAACTGCTGCAACCAAGATTCATGAGACCGGCTGCAAGAGCTGGGTCTGAAC 651
 QY 697 ATGAGCAATTCATTTGATGTCATCTGTGATGATGATGATGATGATGATGATGAT 756
 DB 652 CAGAGAGCAATTTGATGTCATCTGTGATGATGATGATGATGATGATGATGATGAT 711
 QY 757 GGTATCGGGGGGCAAGAGCTGCAAACTTATTTGCAAACTGGTCCATGAAGATC 816
 DB 712 GGCATTGCGGCAAGGCGGTGTCATTCATCAGAAACATGAAGATGAAGATC 771
 QY 817 TTGAGAGATCTTAAATAAGACAGATATCAATTCCTGAGAGCTGCTTACCAAGATC 876
 DB 772 GTGAGGCGGTGAGCCCAAGCAAGTACCCGTTCCAGAGAACTGCTCCACAGAGAGCC 831

QY 877 CGAGCTTGTTCAGAGGCTTAA---TGTACANTGATATTCCTGAGCTAAATGAGACT 933
 DB 832 CAGAGCTTGTTCCTGAGGACAGAGTATGACCCAGAGTCTGTGAGCTGAAGTGAAGC 891
 QY 934 GCACCTGATGAGGAGGCTCATAGTTTCTGTAAAGATTAATGTTTCAACGAAGAT 993
 DB 892 GAGCCAAATGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
 QY 994 CGGGTCAAAAGGCTTAAAGAGATCAATTCGCAAGAAATTAATGTCGCAAGAGAG 1053
 DB 952 CGAATTCGAGTGGGTGCAAGCGCTGAGTAAGAGCCCGGAGGACACCAAGAGAGCT 1011
 QY 1054 CTGAGTCTTCTTTCAGCCCACTGCAACACATGAGCAGCGCTTAAACGAGAGACT 1113
 DB 1012 CTGATGATTTCTTCAAGGAGAGGCTCATCTCTCAGC-----TAAAGCGC 1059
 QY 1114 TCGATTAACCAAGCAAGGCTGCAAGCAAGAAACAAAGCGTGTGGAAGAGAA 1173
 DB 1060 AAGGAGCCAAAGACCAAGGCGCTGCTAAGAAAGAAAGAAAGACTGGGAGCGGAG 1119
 QY 1174 T 1174
 DB 1120 T 1120

RESULT 10
 AAX02109
 ID AAX02109 standard; cDNA; 1149 BP.
 XX
 AC AAX02109;
 XX
 DT 23-APR-1999 (first entry)
 XX
 DE Yeast FEN-1 cDNA.
 XX
 KW FEN-1, yeast; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasia; antineoplastic agent; cleavage; ss.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN US5874283-A.
 XX
 PD 23-FEB-1999.
 XX
 PE 30-MAY-1995; 95US-0455968.
 XX
 PR 30-MAY-1995; 95US-0455968.
 XX
 PA (HARR/) HARRINGTON J J.
 PA (HSIEH/) HSIEH C.
 PA (LIEBER/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX
 DR WPI: 1999-179985/15.
 DR P-PSDB: AAM92506.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 XX e.g. recombinant polypeptides
 XX
 PS Disclosure: Fig 3b; 58pp; English.
 XX
 CC This sequence encodes a yeast FEN-1 (flap endonuclease) protein. This
 CC protein is used in a method to isolate novel human FEN-1 proteins for
 CC detecting a pathological condition in a patient, for diagnostic purposes,
 CC for screening for antineoplastic agents and carcinogens, for diagnostic
 CC staging of neoplasia, for producing recombinant flap endonuclease for use
 CC as research or diagnostic reagents, for producing antibodies reactive
 CC with the novel polypeptides, for producing transgenic nonhuman animals
 CC expressing the novel polypeptides encoded by a transgene. The invention
 CC also provides novel molecular cloning techniques and reagents involving
 CC cleavage of a flap or nick with a flap endonuclease.
 XX


```

Qy 298 GCAAGCCTCCTGATATGAGAAACAAGAGCTGTGTAAGATCTCAAAAAGAGATGAT 357
Db 1321 GGAAGCCACCAATCTCAAGTCCGGTGAAGCTGCCAAGCGCCGAGGGGGAGGAA 1380
Qy 358 GCAACCAAGATCTGACTGAGGAGTAGAGTAGAGATTAAGATGCAATGAAAAATTG 417
Db 1381 GCGGAGAGAGGACTGAAGGCGGACCGATGCGGAGATGATCCGGTATCGAAAAATT 1440
Qy 418 AGCAAGAGAGACTGTAAGAGTACACAGGACACAGAGATGTAAAGGGTATTATAGA 477
Db 1441 AATGCCGATTTGGTCCGGGTACCAAGAGACGCCAAGAGGCCAAGAACTGCTCACA 1500
Qy 478 CTATGGGGGCTCTGTTGTAGAGGACCTTGTGAAGCAAGAACAGATGTCCAGCCCT 537
Db 1501 CTAATGGGTGTGCGCTATGTGTGATGCACCGTGCAGAGCCGAGCCAGTGTCCGCTG 1560
Qy 538 TGCATTAACGATTAAGGTGTGCGGTGTGCTGAGAAATTAAGATGAGACCTTTCTTTGG 597
Db 1561 GTGAAGCTGGAAGGTTTATGCCACCGCCAGGAGATATGGATCCCTCAATTCCGA 1620
Qy 598 GCTCCACGGTCTCTGCTCATTTAATGATCAAGTTCCAGAAATACCTGTGATGAA 657
Db 1621 TCTACAAAGCTGTGAGATACCTTACTACAGGAGGACAGAAAGATGCCGTCAGAG 1680
Qy 658 TTTGATGTGGCAGGTTTGGAGAGCTTGAATCACCATGACCATTTGATTTG 717
Db 1681 TTGAGCTACGACAGCTGTGGAAGGTGTGCGCATTAACATCGAGATTGATCTA 1740
Qy 718 TGCATCTGTGTGATGTGACTATTTGTATAGCATCAAGGATTCGCGGGGCAACAGCT 777
Db 1741 TGTATCTGCGGGTGTGCGATTAAGTGTGAGAGCATCAAGGATTTGAGCCCAAGGAG 1800
Qy 778 CTGAAGCTTATGCTCAACATGGTCCATAGTAAGATCACTCTGTGAGATTTAATAGC 837
Db 1801 ATGCACTGATCAACACCTATCGGATATAGAGACTATCTGTGATACCTGGAGCTTAC 1860
Qy 838 AGATATCAAAATCTCTGAGAGAGCTGCCCTTACCAAGAGCTGAGCTTTCAAGAGCT 897
Db 1861 AAATACACCGTGGCCGAGAACTGGAAGTGGCGGGGAACTCTTCAATCGAACG 1920
Qy 898 AATGTAC--ATTGATATTTCCGAGCTTAATGAGTGCAGCTGACCTGATGAGAGGCTTC 954
Db 1921 GAGTACGTGATGCCGACTCCATAGATCTCAATGGGTGAGCTCGGATGAGAGGCTT 1980
Qy 955 ATAAATTTCTGTAAAGATTAATGTTTCAACGAAGATCGGTTGCAAAAGGCCATAGAG 1014
Db 1981 GTCAAGTTTCTGTGCGGAGACCGGACAGTTCACGAAGAGCGGCTTGGCAACGGTGC 2040
Qy 1015 AAGATCAAAATCTGCCAGATTAATGTCGCAAGGAAGAGTCCGATTTTCAAGGCA 1074
Db 2041 AAGCTGATGAATCCAAAGCAGGCCGACACTCAGGTGAGCTGATAGCTTTTAAAGCA 2100
Qy 1075 ACTGCCACAC 1085
Db 2101 CTGCCACAGC 2111

```

RESULT 12

```

ID ABL10094 standard; cDNA: 5471 BP.
XX ABL10094:
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 24764.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.

```

```

XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PMD, Myers EM;
XX WPI; 2001-656860/75.
XX P-PDB; ABB6591.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Claim 1: SEQ ID NO 24764; 21bp + sequence listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB101840-AB16175), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB101840-AB16175).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
SQ Sequence 5471 BP; 1422 A; 1223 C; 1325 G; 1501 T; 0 other:
Query Match 12.5%; Score 173.2; DB 23; Length 5471;
Best Local Similarity 54.0%; Pred. No. 1.6e-39;
Matches 377; Conservative 0; Mismatches 318; Indels 3; Gaps 1;
Qy 391 GGAATTAAGATCGATGTAAGAAATGAGACAGAGACTGTAAAGTCAAGGCAACAC 450
Db 4 GGAATATATCCCGTATCGAAAGTTAATCCGATTTGTCGGGTACAGAGAGACAC 63
Qy 451 AACGATTTGTAAGAGCTATTAGACTTATGAGGGTCTGTTGTAGAGACACTTCT 510
Db 64 GCCAAGAGGCCAAGAGAACTGCTACACTAATGAGGTGCTTATGATGATGACCGTGC 123
Qy 511 GAAGCAGACAGATGATGAGCGCTTTCATTAAGATGAAGTGTGCGTGTGCTTCA 570
Db 124 GAAAGGAGGCCAGTGTGCTGCTGTGGAAGAGCTGGAAGAGTTTATGCCACCGCAG 183
Qy 571 GAAGATATGAGCTCCCTTACTTTTGGGCTCCAGGGTTCCTTGTCAATTAATGATCA 630
Db 184 GAGATATGATGATCCCTCATATGATGATGATGATGATGATGATGATGATGATGATG 243
Qy 631 AGTTCCAGAAATACCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 690
Db 244 GAGGACAGAAAGATGCCCGTCAAGAGATTCAGTACAGACAGCTGTGGAAGTGTGCC 303
Qy 691 CTCACATGAGCAGTTCATTTGATTTGATGATGATGATGATGATGATGATGATGATG 750
Db 304 ATTAACATATGAGAGTTCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATG 363
Qy 751 ATCAAGAGTATGAGGGGSCAAGACGCTGGAAGCTTATTCGCAACATGAGGATGAG 810
Db 364 ATCAAGGATATGAGACCAAGGAGGATGAGTATGATGATGATGATGATGATGATGATG 423
Qy 811 AGCATCTTGAGAGATCTAATAAAGACAGATATCAAAATCTGAGAGCTGCTTACCAA 870
Db 424 ACTATTTCTGATTAACCTGACTGTACAAATACACCGTCCGAGAACTGGAACATCAAG 483
Qy 871 GAAGCTGAGCGCTTGTCAAGAGGCTATGTCACATTTGATTTGCT--GAGTAAA 927

```

Db 484 GTGGCGGGGAAGCTTTCATCGAACGGAGTAGCTGATGCCAGCTCCATGATGATCTCAA 543
 QY 928 TTGACATGCACCTGATAGAGGGTCTCATTAATCTTCCTGGTAAGATAATAGTTTCAC 987
 Db 544 TGGGTGAGCGCGATAGAGGGCTTGTCAATGTTCTGTGGCGACGGCAGTTCAC 603
 QY 988 GAAGATCGGGTGCACAGGCCATAGAGAGATCAATCTGCCAAGATMAATTCGTGCA 1047
 Db 604 GAAGAGCGGCTTCGCAACGGTCCAAAGAGCTGATGAATCCAAACGAGCCAGACTCAG 663
 QY 1048 GGAAGACTGAGTCTTTTCAAGGCCAATGCGACAC 1085
 Db 664 GTGAGCTGATAGCTTCTTAAAGACTGCCAGCAC 701

RESULT 13
 AAT76685
 ID AAT76685 standard: DNA; 1023 BP.
 AC AAT76685.

DT 14-APR-1998 (first entry)
 DE Pyrococcus furiosus FEN-1 endonuclease gene coding sequence.
 XX

KM Nucleic acid cleavage: 5' nuclease; DNA cleavage: RNA cleavage;
 KM Invader directed cleavage; FEN-1; endonuclease; ds.

OS Pyrococcus furiosus.

PN WO9727214-A1.

PD 31-JUL-1997.

XX 22-JAN-1997; 97WO-US01072.

PR 02-DEC-1996; 96US-0759038.

PR 24-JAN-1996; 96US-059491.

PR 12-JUL-1996; 96US-0682853.

PR 29-NOV-1996; 96US-0756386.

PR 02-DEC-1996; 96US-0758314.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

PI Brow MAD, Dahlberg JE, Hall JG, Kaiser KM, Lyamichiev VI;
 PI Olive DM, Prudent JR;

XX WPI; 1997-393613/36.

DR P-PsDB; AAM24216.

XX Thermostable structure-specific nuclease(s) - used for detection and
 PT characterisation of nucleic acid sequences and variations in nucleic
 PT acid sequences

XX Example 28; Page 283-285; 457pp; English.

XX This sequence comprises the coding region of the gene encoding
 CC Pyrococcus furiosus (Pfu) FEN-1 endonuclease (see AAM24216).
 CC It was obtained by PCR amplification (see AAT76682-83). Large
 CC scale preparation of recombinant Pfu FEN-1 was performed using
 CC E. coli as host. Pfu FEN-1 is a thermostable enzyme. It can be
 CC used in novel methods for the detection and characterisation of
 CC nucleic acid sequences and variations in nucleic acid sequences.

XX Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other;

QY Query Match 11.6%; Score 159.6; DB 18; Length 1023;
 Best Local Similarity 54.6%; Pred. No. 5.9e-36;
 Matches 318; Conservative 0; Mismatches 264; Indels 0; Gaps 0;

QY 86 CGATGAGAGCAGAGATTGCGAGCTACTTGGCGCAAAATCGCGTCGACGCCACAGA 145
 Db 29 CAAGAAAGAAATTGATGAGAAACCTATACGGGAAAGAAATCCATCGACGCTCTTA 88

QY 146 TGAGCATATACAGTTCCTGATTTGATTGGAAAGACAGCGATGAAACTTCACAAATG 205
 Db 89 ATGCATCTACCAATTTTGTCCACATTAAGACAGAAAGATGAACTCCACTATGATTT 148
 QY 206 AAGCTGTGCAAGTCACTAGTATTTGCAAGGAATGTTCAACCGGACAAATAGATTACTG 265
 Db 149 CAAAGGGTGAATTAACCTCCACCTTAAGCGGGCTCTTTACAGSACAATTAACCTAATG 208
 QY 266 AAGCGGGAATCAAGCCACTTATGTTTGTGCAAGCCCTCCCTGATATGAAGAAACAG 325
 Db 209 AGCTGGGAATTAACCTGATGTTTGTGCAAGCCCTCCCTGATATGAAGAAACAG 268
 QY 326 AGCTTGTAAAGATGCTCAAAAGAGATGATGCAACCAAGATCTGACTGAGGCAATG 385
 Db 269 AGCTGCAAAAG 328
 QY 386 AGGTGAGATTAAGATGCAATGTAAGAAATGACCAAGAGAGAGAGAGAGAGAGAGAG 445
 Db 329 AAAAGAGAGATGAG 388
 QY 446 AACACAGCAGAGATTGTAAGCGCTATTAGACTTATGAGGGGTTCTGTTGAGAGGAC 505
 Db 389 TCGTCTATCGAGAGTCAAAAGAACTTGAAGCTTATGAGGATTCATAGTTCAACAC 448
 QY 506 CTTTGAAGCAGACAGATGTCAGGCCCTTTCATTAAGATTAAGGTTGCTGTTG 565
 Db 449 CTAGCGAGGAG 508
 QY 566 CTTGAGAGATATGAGTCCCTTACTTTTGGGCTCCAGCGTTCCTGCTGATTAATG 625
 Db 509 CTAGTCAAGATTAGCATTTCTTACTTTTGGAGCTCCAGACTGTTAAGAACTTAACA 568
 QY 626 ATCCAGTTCAGAAATTAACCTGATGCAATTTGATGTTG 667
 Db 569 TAAACGAGAAAGAAAGTTGCTGGGAAATGTCATGCTG 610

RESULT 14
 AAV65840
 ID AAV65840 standard: DNA; 1023 BP.
 XX

AC AAV65840;

DT 02-FEB-1999 (first entry)

DE Pyrococcus furiosus FEN-1 endonuclease gene ORF.

KM Nucleic acid detection; multiple sequential invasive cleavage;
 KM FEN-1; endonuclease; nuclease; ds.

XX Pyrococcus furiosus.

OS WO9842873-A1.

PN 01-OCT-1998.

PD 24-MAR-1998; 98WO-US05809.

PR 24-MAR-1997; 97US-0823516.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

PI Brow MAD, Hall JG, Kwiatkowski RW, Lyamichiev VI;
 PI Mast AL, Vavra SH;

XX WPI; 1998-557036/47.

DR P-PsDB; AAM79970.

XX Detecting target nucleic acid by sequence-specific cleavage of
 PT complex with two specific oligonucleotides - used to detect
 PT cytomegalovirus DNA

XX

Tue May 27 16:02:22 2003

us-09-805-311-5.std.rng

Page 15

OY	326	AGCTGCTTAAAGATACCTCAAAAAGCATATGTCACCAAAATCTGACTAGCACTAG	385
Db	269	AGCTCGAAAAAAGAGAGACACGACAGAGCACTGACAGAAAAGTGGAGACAACTTGG	328
OY	386	AGGTAGAGATTAAGATCCGATTAAAAATTGACGAAGAGACTGTAAGCTCACAAGGC	445
Db	329	AAAAAGAGAGATTAAGAGAACCAAGAAAATATGCCCAAAAGACCAACAGGGTAAATGAA	388
OY	446	AACACACGACAAATTTGTAAAGCGCTATTAAAGCTATTGAGGGCTTCTCTTTAGAGCAC	505
Db	389	TGCTCATCGAGGTGCAAAAAAAGCTTTAGAGCTTATGGGAATTCCTTATGTCACAC	448
OY	506	CTTCTGACAGCAGACAGATATGTGCAGCCCTTGCATAAACGATTAAGTGTTCGCTGTG	565
Db	449	CTAGCGGAGAGAGGCCCAAGCTGCATATATGCGCCCAAAAGAGAGCGGTATGCTACGG	508
OY	566	CTTCAGAGATATGCACTCCCTACTTTTGGGGCTCCACGGTTCCTGTCATTTAAATGG	625
Db	509	CTAGTCAGATTCACGATTCCTACTTTTTTGGAGCTCCAGACACTGTGAAAACTTAACAA	568
OY	626	ATCCAGATTCGAAGAAAATTAACCTGTGATGAGCAATTTGATGTG	667
Db	569	TTACAGAAAAAGCAAGTTCCTGGGAAAAAATGTCTTACGTCG	610

Search completed: November 5, 2002, 11:01:08
Job time : 179.876 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:33:28 ; Search time 33.87 seconds
(without alignments)
10015.359 Million cell updates/sec

Title: US-09-805-311-5

Perfect score: 1381

Sequence: 1 cgaccacgcgcgcgcgcac.....ttgaaaaaaaaaaaaa 1381

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 38353 seqs, 122816752 residues.

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA: *

1: /cgn2.6/prodata/1/lna/5A.COMB.seq:*

2: /cgn2.6/prodata/1/lna/5B.COMB.seq:*

3: /cgn2.6/prodata/1/lna/6A.COMB.seq:*

4: /cgn2.6/prodata/1/lna/6B.COMB.seq:*

5: /cgn2.6/prodata/1/lna/PCFUS.COMB.seq:*

6: /cgn2.6/prodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1381	100.0	1381	4	US-09-426-557-5
2	1350.2	97.8	1463	4	US-09-426-557-1
3	1348.6	97.7	1541	4	US-09-426-557-3
4	1334.2	96.6	1478	4	US-09-426-557-7
5	302.6	21.9	1144	2	US-08-455-968E-2
6	302.6	21.9	1144	2	US-08-455-968E-28
7	276.6	20.0	1033	2	US-08-455-968E-9
8	272.2	19.7	1930	2	US-08-455-968E-4
9	270.2	19.6	1149	2	US-08-455-968E-6
10	158.6	11.6	1023	2	US-08-757-653-175
11	158.6	11.6	1023	2	US-08-823-516-78
12	158.6	11.6	1023	3	US-08-759-038-114
13	135.6	11.6	1023	3	US-08-758-314-114
14	135.6	9.8	1032	4	US-09-146-319-1
15	135.6	9.8	1032	4	US-09-175-973-1
16	89.2	6.5	981	2	US-08-757-653-171
17	89.2	6.5	981	2	US-08-823-516-74
18	89.2	6.5	981	3	US-08-759-038-110
19	89.2	6.5	981	3	US-08-758-314-110
20	65.2	4.7	7218	1	US-08-232-463-14
21	47	3.4	455	1	US-08-636-928-4
22	37.6	2.7	289	4	US-09-007-005-17
23	37.6	2.7	289	4	US-09-244-796-17
24	37.6	2.7	2339	4	US-09-268-140-11
25	37.6	2.7	2505	4	US-09-268-140-11
26	37.6	2.7	2517	4	US-09-268-140-7
27	36.4	2.6	1412	1	US-08-097-831-1

28	34.4	2.5	3825	4	US-09-208-742-3	Sequence 3, Appl1
29	34.2	2.5	5361	4	US-08-973-462-2	Sequence 2, Appl1
30	34.2	2.5	6152	4	US-08-973-462-1	Sequence 1, Appl1
31	34	2.5	677	3	US-08-881-094-15	Sequence 15, Appl1
32	34	2.5	36741	4	US-09-301-665-3	Sequence 3, Appl1
33	33.8	2.4	3791	4	US-09-377-310-1	Sequence 1, Appl1
34	33.8	2.4	43804	4	US-09-171-461-1	Sequence 1, Appl1
35	33.6	2.4	1272	2	US-08-972-258-1	Sequence 1, Appl1
36	33.6	2.4	1272	4	US-09-263-128-1	Sequence 1, Appl1
37	33.6	2.4	2346	4	US-09-193-503B-4	Sequence 4, Appl1
38	33.6	2.4	2346	4	US-09-193-503B-5	Sequence 5, Appl1
39	33.6	2.4	2346	4	US-09-193-503B-8	Sequence 8, Appl1
40	33.6	2.4	72604	4	US-09-268-992-7	Sequence 7, Appl1
41	33.4	2.4	984	4	US-08-927-219-52	Sequence 52, Appl1
42	33.4	2.4	1858	3	US-08-742-185-96	Sequence 96, Appl1
43	33.4	2.4	10014	4	US-08-927-219-130	Sequence 130, Appl1
44	33.2	2.4	48974	4	US-08-920-422-17	Sequence 17, Appl1
45	33	2.4	23673	4	US-09-773-816-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
US-09-426-557-5
; Sequence 5, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; EARLIER FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: CDS
; LOCATION: (37)...(1173)
US-09-426-557-5
```

```
Query Match          100.0%; Score 1381; DB 4; Length 1381;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	CGACCCAGCGCGTCCGCCACAGCCGCCGACAGATGGATTCAGGATTTGACGANA	60
DB	1	CGACCCAGCGCGTCCGCCACAGCCGCCGACAGATGGATTCAGGATTTGACGANA	60
QY	61	CTGCTGCGCGGACATGCGCCCAAGCGCATGAAGAGAGAGTTCAGAGCTTCTGGC	120
DB	61	CTGCTGCGCGGACATGCGCCCAAGCGCATGAAGAGAGAGTTCAGAGCTTCTGGC	120
QY	121	CGCAAAATCGCGTCCGCCACAGCCGCCGATATATTCATTTCTAGTTGGAAG	180
DB	121	CGCAAAATCGCGTCCGCCACAGCCGCCGATATATTCATTTCTAGTTGGAAG	180
QY	181	ACAGGATGGAATCTCTCAATATGAAGCTGTAACTCTAGTATTTGCAAGAA	240
DB	181	ACAGGATGGAATCTCTCAATATGAAGCTGTAACTCTAGTATTTGCAAGAA	240
QY	241	TTCAACCGGACATTAAGTTACTGGAACGGGAATCAAGCCAGTTATGTTGATG	300
DB	241	TTCAACCGGACATTAAGTTACTGGAACGGGAATCAAGCCAGTTATGTTGATG	300
QY	301	AAGCCCTGATATGAAGAACAAGAGCTTCTAAAGTACTCAAAAAGAGATGATGA	360
DB	301	AAGCCCTGATATGAAGAACAAGAGCTTCTAAAGTACTCAAAAAGAGATGATGA	360

Db 301 AAGCTCTGATATGAGAACAGAGCTTGCTAAAGATCTCAAAAAGATGATGCA 360
OY 361 ACCAAGATGTGACTGAGCAGTAGAGTAGAGATAAGATGGCATTGAAAATTGAGC 420
Db 361 ACCAAGATGTGACTGAGCAGTAGAGTAGAGATAAGATGGCATTGAAAATTGAGC 420
OY 421 AAGAGAGCTGTAAGGTGCAACAGCAACAGAGATTGTAAAGCGCTATTAAAGCTT 480
Db 421 AAGAGAGCTGTAAGGTGCAACAGCAACAGAGATTGTAAAGCGCTATTAAAGCTT 480
OY 481 ATGGGGTTCCTGTTAGAGCACCCTTCTGAGAGAGAGAGATGTCAGCCCTTTTC 540
Db 481 ATGGGGTTCCTGTTAGAGCACCCTTCTGAGAGAGAGAGATGTCAGCCCTTTTC 540
OY 541 ATTAACATAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 ATTAACATAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
OY 601 CCACGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CCACGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
OY 661 GATGTTCCAGAGTTTGGAGAGAGCTGTAACCTGACCATGACCGATTGATTGCTG 720
Db 661 GATGTTCCAGAGTTTGGAGAGAGCTGTAACCTGACCATGACCGATTGATTGCTG 720
OY 721 ATCTGTGTGATGTGACTGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 ATCTGTGTGATGTGACTGATGATGATGATGATGATGATGATGATGATGATGATG 780
OY 781 AAATTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 781 AAATTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
OY 841 TATTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 841 TATTAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
OY 901 GTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 901 GTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
OY 961 TTCTGTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 961 TTCTGTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
OY 1021 AAATCTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 AAATCTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
OY 1081 ACCACATCAGCAGCGCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 ACCACATCAGCAGCGCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
OY 1141 AACCAAGAAACAAAGCGTGTGGAAGAAATATATCTGATGCTGTTGATGCAACTA 1200
Db 1141 AACCAAGAAACAAAGCGTGTGGAAGAAATATATCTGATGCTGTTGATGCAACTA 1200
OY 1201 CGACTAGCAAGAGAGCGGTGCGTGTGATGATGATGATGATGATGATGATGATG 1260
Db 1201 CGACTAGCAAGAGAGCGGTGCGTGTGATGATGATGATGATGATGATGATGATG 1260
OY 1261 ACTGAGAGCTGTTGTTAAAGTTGCGCCATGTTCAAGCTGGGGTAAAGTGTGTTG 1320
Db 1261 ACTGAGAGCTGTTGTTAAAGTTGCGCCATGTTCAAGCTGGGGTAAAGTGTGTTG 1320
OY 1321 AAGAGATTGTTGTTAAAGTTGCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1380
Db 1321 AAGAGATTGTTGTTAAAGTTGCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1380
OY 1381 A 1381
Db 1381 A 1381

RESULT 2
US-09-426-557-1
; Sequence 1, Application US/09426557
; Patient No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85) ... (1221)
US-09-426-557-1
Query Match 97.8%, Score 1350.2; DB 4; Length 1463;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 CGACCCAGCGTCGCGCCAGACGCCCGCAGAGATGAGGATCAAGGGTTTGACGAA 60
Db 49 CGCGCGCGCCAGACCGCCAGACGCCCGCAGAGATGAGGATCAAGGGTTTGACGAA 108
OY 61 CTCTGCGCGGACATGCGCCCAAGGCGATGAAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 109 CTCTGCGCGGACATGCGCCCAAGGCGATGAAGAGAGAGAGAGAGAGAGAGAGAG 168
OY 121 CGCAAAATCGCGTGCAGCCGACATGAGATACCATGATGATGATGATGATGATG 180
Db 169 CGCAAAATCGCGTGCAGCCGACATGAGATACCATGATGATGATGATGATGATG 228
OY 181 ACAGGATGAGAACTGTCACAAATGAAGTGTGATGATGATGATGATGATGATG 240
Db 229 ACAGGATGAGAACTGTCACAAATGAAGTGTGATGATGATGATGATGATGATG 288
OY 241 TTCACCGGACATTAAGATTACTGAGAGCGGAAATCAAGCAGTTTATGTTGATG 300
Db 289 TTCACCGGACATTAAGATTACTGAGAGCGGAAATCAAGCAGTTTATGTTGATG 348
OY 301 AAGCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 349 AAGCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 408
OY 361 ACCAAGATGCTGACTGAGGAGTAGAGTAGAGATGAAGTGGGATGAAAATTGAG 420
Db 409 ACCAAGATGCTGACTGAGGAGTAGAGTAGAGATGAAGTGGGATGAAAATTGAG 468
OY 421 AAGAGAGCTGTAAGGTGCAACAGCAACAGAGATTGTAAAGCGCTATTAAAGCTT 480
Db 469 AAGAGAGCTGTAAGGTGCAACAGCAACAGAGATTGTAAAGCGCTATTAAAGCTT 528
OY 481 ATGGGGTTCCTGTTAGAGCACCCTTCTGAGAGAGAGAGATGTCAGCCCTTTTC 540
Db 529 ATGGGGTTCCTGTTAGAGCACCCTTCTGAGAGAGAGAGATGTCAGCCCTTTTC 588
OY 541 ATTAACATAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 589 ATTAACATAGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
OY 601 CCACGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 649 CCACGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708

QY	661	GATGTGGCCAAAGCTTTTGGAGACCTTGAACATACCATGGACCACTGATGATTTGCTG	720
Dp	709	GATGTGGCCAAAGCTTTTGGAGACCTTGAACATACCATGGACCACTGATGATTTGCTG	768
QY	721	ATCTGTGTGATGTGACTATTGTGATAGCATCAAAAGTATCGGGGGCAACAGCTCTG	780
Dp	769	ATCTGTGTGATGTGACTATTGTGATAGCATCAAAAGTATCGGGGGCAACAGCTCTG	828
QY	781	AACTATTTCGCAACATGAGGCTCCATGGAAGCATCTTGTGGAACTTAAATAAGACAGA	840
Dp	829	AACTATTTCGCAACATGAGGCTCCATGGAAGCATCTTGTGGAACTTAAATAAGACAGA	888
QY	841	TATCAAAATTCCTGAGAGACTGGCTTACCAAGAGCTCGACGCTTGTCTAAGSACCTAAT	900
Dp	889	TATCAAAATTCCTGAGAGACTGGCTTACCAAGAGAGCTCGACGCTTGTCTAAGSACCTAAT	948
QY	901	GTCACATTTGGATATTCCTCGAGCTTAAATTGGACTGCACCTGATGAGAGGCTCTATAGT	960
Dp	949	GTCACATTTGGATATTCCTCGAGCTTAAATTGGACTGCACCTGATGAGAGGCTCTATAGT	1000
QY	961	TTCTGTGTTAAAGTAATAGTTTTCACAGAGATCGGGGTACCAAGGCCATFGAGAAATC	1020
Dp	1009	TTCTGTGTTAAAGTAATAGTTTTCACAGAGATCGGGGTACCAAGGCCATFGAGAAATC	1060
QY	1021	AAATCTGCCAGATTAATTCGTGCCAAGGAAGCTGAGTCTCTTTTCAAGCCAACTGCC	1080
Dp	1069	AAATCTGCCAGATTAATTCGTGCCAAGGAAGCTGAGTCTCTTTTCAAGCCAACTGCC	1120
QY	1081	ACCACATTCGACACCGCTTAAAGCGGAAGACACTTCGATTAACCAAGCAAGGACGTCGC	1140
Dp	1129	ACCACATTCGACACCGCTTAAAGCGGAAGACACTTCGATTAACCAAGCAAGGACGTCGC	1188
QY	1141	AACAATAAAACAAGCGGTGGTGGAAATAAATAATATCTGATGCTGTGATGTACACTA	1200
Dp	1189	AACAATAAAACAAGCGGTGGTGGAAATAAATAATATCTGATGCTGTGATGTACACTA	1240
QY	1201	CGACTCGAAGACAGCGGTGGCGTGATCACTTCGCTTAGATTAATTTAACTCCCTGTTTTA	1260
Dp	1249	CGACTCGAAGACAGCGGTGGCGTGATCACTTCGCTTAGATTAATTTAACTCCCTGTTTTA	1300
QY	1261	ACACAGAGCTTTGGTAAAGTATCCCATCTTTCAAGCTGGGGTAAAGTAGTGTGTTTTG	1320
Dp	1309	ACACAGAGCTTTGGTAAAGTATCCCATCTTTCAAGCTGGGGTAAAGTAGTGTGTTTTG	1360
QY	1321	AACAGATTGGGTACCAAGTAAACAAACTTATCGCTGTTTTTTT	1363
Dp	1369	AACAGATTGGGTACCAAGTAAACAAACTTATCGCTGTTTTTTT	1411

RESULT 3
 US-09-426-557-3
 Sequence 3, Application US/09426557
 Patent No. 6232527
 GENERAL INFORMATION:
 APPLICANT: Mahajan, Pramod B.
 TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
 TITLE OF INVENTION: Thereof
 FILE REFERENCE: 0961
 CURRENT APPLICATION NUMBER: US/09/426,557
 CURRENT FILING DATE: 1999-10-22
 EARLIER APPLICATION NUMBER: 60/112,332
 EARLIER FILING DATE: 1998-12-15
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 3
 LENGTH: 1541
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (79) ... (1215)
 US-09-426-557-3

Query Match	97.7%	Score 1348.6;	DB 4;	Length 1541;
Best Local Similarity	99.3%;	Pred. No. 0;		
Matches 1354; Conservative	0;	Mismatches	9;	Indels 0; Gaps 0;

QY	1	CGACCAAGGCTGCCGGGCGACAGCCCGGAGACGAGATGGGCATTCAGAGGGTTTGAGCAAA	60
Db	43	CGCGGCCCGCCACCCGCCACACAGCCCGCGACAGAGATGGGCATTCAGAGGGTTTGAGCAAA	102
QY	61	CTGCTGGCGGACAAATGCGGCCCAAGCGGATGAGAGAGCAAGATTGAGAGCTACTTCGGC	120
Db	103	CTGTGGCGGACAAATGCGGCCCAAGCGGATGAGAGAGCAAGATTGAGAGCTACTTCGGC	162
QY	121	CGCAAAATGCGCGTGGAGCGCCAGCATGAGCATATCCGATTCCGATTGTGATTTGGGAAGC	180
Db	163	CGCAAAATGCGCGTGGAGCGCCAGCATGAGCATATCCGATTCCGATTGTGATTTGGGAAGC	222
QY	181	ACAGGCTATGGAACCTCTCACAATGAAAGCTGGTGAAGTCACTAGTCAATTTGCCAAGGAATG	240
Db	223	ACAGGCTATGGAACCTCTCACAATGAAAGCTGGTGAAGTCACTAGTCAATTTGCCAAGGAATG	282
QY	241	TTTCAACCGGACAAATAGATTTACTGGGAAGCGGGAATCAAGCCAGTTTATGTTTTGATGGGC	300
Db	283	TTTCAACCGGACAAATAGATTTACTGGGAAGCGGGAATCAAGCCAGTTTATGTTTTGATGGGC	342
QY	301	AAGGCTCTGATATGAGAAACAGAGGCTTGTAAAGATACTGCAAAAAGAGATGATGCA	360
Db	343	AAGGCTCTGATATGAGAAACAGAGGCTTGTAAAGATACTGCAAAAAGAGATGATGCA	402
QY	361	ACCAAAATGTCAGTCAAGCAGTAAAGAGTAAAGATTTGGAATTTGAATTTGAGC	420
Db	403	ACCAAAATGTCAGTCAAGCAGTAAAGAGTAAAGATTTGGAATTTGGAATTTGAGC	462
QY	421	AAGAGACTGTAAAGGTTCACAAGGCAACACAACGAAATTTGTAAACGGCTATTAAAGCTT	480
Db	463	AAGAGACTGTAAAGGTTCACAAGGCAACACAACGAAATTTGTAAACGGCTATTAAAGCTT	522
QY	481	ATGGGGGTTCTCTTTTGAAGAGCACTTCTGAAAGCAAGAAATGTCAGCCCTTTGCG	540
Db	523	ATGGGGGTTCTCTTTTGAAGAGCACTTCTGAAAGCAAGAAATGTCAGCCCTTTGCG	582
QY	541	ATAAAGCATTAAGGTGTTCGCTTTCGTCAGAAAGATATGAGTCCCTTACTTTTGGGGCT	600
Db	583	ATAAAGCATTAAGGTGTTCGCTTTCGTCAGAAAGATATGAGTCCCTTACTTTTGGGGCT	642
QY	601	CCAGGTTCTTGTGCTATTTAATGGAATCCAACTTCCAAAGAAATATCCGTGATGGAATTT	660
Db	643	CCAGGTTCTTGTGCTATTTAATGGAATCCAACTTCCAAAGAAATATCCGTGATGGAATTT	702
QY	661	GATGTTGCCAAGGTTTGGAGAGGCTTGAACACACCATGAGGACCGATTGATTTATGCG	720
Db	703	GATGTTGCCAAGGTTTGGAGAGGCTTGAACACACCATGAGGACCGATTGATTTATGCG	762
QY	721	ATCTCTGTCTGATGTGACTATTGTGATACCATCAAAAGGTATTCGGGGGCAACAGCTCTG	780
Db	763	ATCTCTGTCTGATGTGACTATTGTGATACCATCAAAAGGTATTCGGGGGCAACAGCTCTG	822
QY	781	AAACTTATTTGTTGATGAGTGGTCCATATGAAAGCATCTTGGAGAAATCTTAATAAAGACAGA	840
Db	823	AAACTTATTTGTTGATGAGTGGTCCATATGAAAGCATCTTGGAGAAATCTTAATAAAGACAGA	882
QY	841	TATCAAAATTCCTGAGGACTGGCTTACCAAGAAGCTCGAGCTTGTTAAGGAGCTTAAT	900
Db	883	TATCAAAATTCCTGAGGACTGGCTTACCAAGAAGCTCGAGCTTGTTAAGGAGCTTAAT	942
QY	901	GTCACATTTGATATTCCTGAGTAAATATGAGTGCACATGATGAGGAGGGCTCATTAAGT	960
Db	943	GTCACATTTGATATTCCTGAGTAAATATGAGTGCACATGATGAGGAGGGCTCATTAAGT	1002
QY	961	TTTCTGTTGTTAAAGATATATGTTTTCACAGAAATCGGGTGTACAAAGGCTATGAGAAATC	1020
Db	1003	TTTCTGTTGTTAAAGATATATGTTTTCACAGAAATCGGGTGTACAAAGGCTATGAGAAATC	1062

```

OY 1021 AAATCTGCCAAGATAAATCTGTCGAAGGAGACTCGATCTCTTTTCAAGCCAACTGCC 1080
    |||||||
Db 1063 AAATCTGCCAAGATAAATCTGTCGAAGGAGACTCGATCTCTTTTCAAGCCAACTGCC 1122
OY 1081 ACCACATCAGCACCGCTAAACGGAAGGAGACTTCGGATTAACAGCAAGGACGCTGGC 1140
    |||||||
Db 1123 ACCACATCAGCACCGCTAAACGGAAGGAGACTTCGGATTAACAGCAAGGACGCTGGC 1182
OY 1141 AACAGAAAAACAAGGCTGGTGGAAAGAAATATCTTGGATCGCTTGGATGTCAACTA 1200
    |||||||
Db 1183 AACAGAAAAACAAGGCTGGTGGAAAGAAATATCTTGGATCGCTTGGATGTCAACTA 1242
OY 1201 CGACTACGAAAGACGCGTGGCGCTGATCACTTCGCTAGATTAATTAACCTCCCTGTTTA 1260
    |||||||
Db 1243 CGACTACGAAAGACGCGTGGCGCTGATCACTTCGCTAGATTAATTAACCTCCCTGTTTA 1302
OY 1261 ACTCAGAGCTTTGGTAAAGTTCCGCCATGTTCAAGCTGGGGGTAAAGTATGTGTGTTG 1320
    |||||||
Db 1303 ACTCAGAGCTTTGGTAAAGTTCCGCCATGTTCAAGCTGGGGGTAAAGTATGTGTGTTG 1362
OY 1321 AAGAGATTGGTGTACCAAGTAAACAAACTTATCGCTGTTTTTT 1363
    |||||||
Db 1363 AAGAGATTGGTGTACCAAGTAAACAAACTTATCGCTGTTTTTT 1405

```

RESULT 4

```

US-09-426-557-7
: Sequence 7, Application US/09426557
: Patent No. 623527
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
: FILE OF INVENTION: Thereof
: FILE REFERENCE: 0961
: CURRENT APPLICATION NUMBER: US/09/426,557
: EARLIER FILING DATE: 1999-10-22
: EARLIER APPLICATION NUMBER: 60/112,332
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 7
: LENGTH: 1478
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (97)...(1233)
US-09-426-557-7

```

```

Query Match          96.6%: Score 1334.2; DB 4; Length 1478;
Best Local Similarity 98.7%: Pred. No. 0; Mismatches 18; Indels 0; Gaps 0;
Matches 1345; Conservative 0;

```

```

OY 1 CGACCCACGCGTCCGGCCACAGCCGCCGACAGAGATGGGCATCAAGGTTTGACGAAA 60
    |||||||
Db 61 CGCGCGCCGCCACCGCCACAGCCGCCGACAGAGATGGGCATCAAGGTTTGACGAAA 120
OY 61 CTGCTGCGGACAAATGGCGCAAGGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
    |||||||
Db 121 CTGCTGCGGACAAATGGCGCAAGGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 160
OY 121 CGCAAAATCGCGCTGACGACGACATGAGCATATACCAATTCCTGATTTAGTTGAAG 180
    |||||||
Db 181 CGCAAAATCGCGCTGACGACGACATGAGCATATACCAATTCCTGATTTAGTTGAAG 240
OY 181 ACAGGCAATGGAACCTCACAATAATGAAGTGTGAAGTCACTAGCATTTGCAAGGATG 240
    |||||||
Db 241 ACAGGCAATGGAACCTCACAATAATGAAGTGTGAAGTCACTAGCATTTGCAAGGATG 300
OY 241 TTCAACCGGACAAATAGATTACTGGAAGCGGGAATCAAGCAAGTTTATGTTTGAATGC 300
    |||||||
Db 301 TTCAACCGGACAAATAGATTACTGGAAGCGGGAATCAAGCAAGTTTATGTTTGAATGC 360

```

```

OY 301 AAGCCTCTGATATGAAGAAACAAGAGCTTGTCTAAAGATCTCAAAAAGAGATGATGCA 360
    |||||||
Db 361 AAGCCTCTGATATGAAGAAACAAGAGACTTGTCTAAAGATCTCAAAAAGAGATGATGCA 420
OY 361 ACCAAGATCTGACTAGAGGAGTAGAGTAGAGATTAAGATTCGATTTGAAAAATTAGC 420
    |||||||
Db 421 ACCAAGATCTGACTAGAGGAGTAGAGTAGAGATTAAGATTCGATTTGAAAAATTAGC 480
OY 421 AAGAGACTGTAAAGGTCAACAAGGACACACAGCAAGATTTGAAACGACTTTAAGACT 480
    |||||||
Db 481 AAGAGACTGTAAAGGTCAACAAGGACACACAGCAAGATTTGAAACGACTTTAAGACT 540
OY 481 ATGGGGTTCCTGTTGTAGAGGACCTTGTGAAGCAAGCAAGATGTGACGCCCTTTGC 540
    |||||||
Db 541 ATGGGGTTCCTGTTGTAGAGGACCTTGTGAAGCAAGCAAGATGTGACGCCCTTTGC 600
OY 541 ATTAACGATTAAGGTGTCGCTGTCCTTCAAGAAATATGAGACTCCCTTACTTTGGGGCT 600
    |||||||
Db 601 ATTAACGATTAAGGTGTCGCTGTCCTTCAAGAAATATGAGACTCCCTTACTTTGGGGCT 660
OY 601 CCACGTTCTCTGTCATTTATGATCCAAAGTCCCAAGAAATACCTGTATGGAATTT 660
    |||||||
Db 661 CCACGTTCTCTGTCATTTATGATCCCAAGTCCCAAGAAATACCTGTATGGAATTT 720
OY 661 GATGTGCAAGGTTTGGAGAGCTTGAACACACAGCAAGCAAGTATGATTTGTC 720
    |||||||
Db 721 GATGTGCAAGGTTTGGAGAGCTTGAACACACAGCAAGTATGATTTGTC 780
OY 721 ATCTGTGTGATGTGACTATTTGTGATACATCAAAAGTATCGGGGGCAACAGCTCTG 780
    |||||||
Db 781 ATCTGTGTGATGTGACTATTTGTGATACATCAAAAGTATCGGGGGCAACAGCTCTG 840
OY 781 AAATTAATGTGTCAACAVGGGTCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGA 840
    |||||||
Db 841 AAATTAATGTGTCAACAVGGGTCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGA 900
OY 841 TATCAAAATTCCTGAGAGCTGCTTACCAAGAGAGCTCGACGCTTGTTCMAAGACCTAAT 900
    |||||||
Db 901 TATCAAAATTCCTGAGAGCTGCTTACCAAGAGAGCTCGACGCTTGTTCMAAGACCTAAT 960
OY 901 GTCAATTTGGATATTCCTGAGCTAAATAATGAGTGCACCTGATGAGAGAGCTCATAGT 960
    |||||||
Db 961 GTCAATTTGGATATTCCTGAGCTAAATAATGAGTGCACCTGATGAGAGAGCTCATAGT 1020
OY 961 TTCTGTGTAAGATTAATGTTTCAACGAAGATCGGGTACAAAGCCATAGAGAGATC 1020
    |||||||
Db 1021 TTCTGTGTAAGATTAATGTTTCAACGAAGATCGGGTACAAAGCCATAGAGAGATC 1080
OY 1021 AAATCTGCCAAGATAAATCTGTCGAAGGAGACTCGATCTCTTTTCAAGCCAACTGCC 1080
    |||||||
Db 1081 AAATCTGCCAAGATAAATCTGTCGAAGGAGACTCGATCTCTTTTCAAGCCAACTGCC 1140
OY 1081 ACCACATCAGCACCGCTAAACGGAAGGAGACTTCGGATTAACAGCAAGGACGCTGGC 1140
    |||||||
Db 1141 ACCACATCAGCACCGCTAAACGGAAGGAGACTTCGGATTAACAGCAAGGACGCTGGC 1200
OY 1141 AACAGAAAAACAAGGCTGGTGGAAAGAAATATCTTGGATCGCTTGGATGTCAACTA 1200
    |||||||
Db 1201 AACAGAAAAACAAGGCTGGTGGAAAGAAATATCTTGGATCGCTTGGATGTCAACTA 1260
OY 1201 CGACTACGAAAGACGCGTGGCGCTGATCACTTCGCTAGATTAATTAACCTCCCTGTTTA 1260
    |||||||
Db 1261 CGACTACGAAAGACGCGTGGCGCTGATCACTTCGCTAGATTAATTAACCTCCCTGTTTA 1320
OY 1261 ACTCAGAGCTTTGGTAAAGTTCCGCCATGTTCAAGCTGGGGGTAAAGTATGTGTGTTG 1320
    |||||||
Db 1321 ACTCAGAGCTTTGGTAAAGTTCCGCCATGTTTCAAGCTGGGGGTAAAGTATGTGTGTTG 1380
OY 1321 AAGAGATTGGTGTACCAAGTAAACAAACTTATCGCTGTTTTTT 1363
    |||||||
Db 1381 AAGAGATTGGTGTACCAAGTAAACAAACTTATCGCTGTTTTTT 1423

```

```

RESULT 5
US-08-455-968E-2
; Sequence 2, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-455-968E-2

Query Match      21.9%  Score 302.6;  DB 2;  Length 1144;
Best Local Similarity 56.6%;  Pred. No. 2.2e+80;
Matches 601;  Conservative 0;  Mismatches 454;  Indels 6;  Gaps 2;

```

```

QY 457 GATTGTAACGCGCTATTAGACTTATGGGGTTCCTGTTGTAGAGCCACCTCTTGAAGA 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 GAGTGCACAAACATCTGCTGACCCATCCATCTGATGATGATGATGATGATGATGATG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 GAACGAGAAATGTGCGAGCTTTGCTTAACGATAGAGTGTGCTGTGCTTCAAGAT 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 GAGGCCACGTGTGCTGTGCTGTGGAAGGCTGCGCAAAAGTCTATGCTGTGCTACCGAGAC 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 577 ATGAGTCCCTTACTTTTGGGGCTCCAGCGTTCCTTGTGATTTATGATGATCAAGTTC 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 ATGAGTCCCTTACTTTGCGACGCTGTGATGATGATGATGATGATGATGATGATGATG 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 637 AAGAAATACCTGTGATGATTTGATTTGATGATGATGATGATGATGATGATGATGATG 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 598 AAAAAGCTCCCAATCCAGAAATTCACCGAGCCGATTCCTCAGAGCTGTGCGCTTGAAC 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 ATGACCACTTCAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 658 CAGGAACAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 GGTATGGGGGGGCAACAGCTTGAACCTTATTCGTAACATGGGTCCATAGAAAGCATC 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 718 GGTATTTGGGCGCCAAAGCGGCTGTGACCTCATCCAGAACCAAGAGCATCGAGAGATC 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 TTGAGAAATCTTAATTAAGACATATCAAAATTCCTGAGAGATGCGCTTACCAAGACT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 778 GTTCGGCGACTTGAACCCCAACAGATACCTGTGCGCAAGAAATTTGCTCCACAAAGAGCT 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 CGAGCGCTTTGCAAGAGCGCTAATG--TCACATTGATATTCCTGAGCTAAATAGACT 933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 838 CACGACCTCTTCTTGTGAACTGTGAGTGTGAGCCAGAGATCTGTGAGCTGAAGTGGAGC 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 934 GCACCTGATGAGAGGCTCATTAAGTTCTTCTGTTAAAGATTAATGTTTCAACGAGAT 993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 898 GAGCCAAATGAAGAAAGATGATCAAGTTCAATGTTGTGTAAGAAACAGATTTCTTGAGAG 957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 994 CGGGTACAAAGCCATAGAGATCAAAATCTGCCAAGATTAATGCTGCAAGAAAGA 1053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 958 CGAATCCGCAATGGGTCAAGAGGCTGAGTAAAGAGCCGCCAAGAGCAAGCAACCCAGGGCCGC 1017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1054 CTGAGTCTCTTTTCAAGCAACTGCCACACATGACGAC 1094
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1018 CTGAGATGATTTCTTCAAGGTGACCGCTCATCTCTTGAGC 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-08-455-968E-28
; Sequence 28, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

```

NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (polynucleotide)
US-08-455-968E-28

Query Match
Best local Similarity 56.6%; Score 302.6; DB 2; Length 1144;
Pred. No. 2.2e-80;
Matches 601; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 37 ATGGGCAATCAAGGCTTGGAGAACTGCGCGGACATGCGCCCAAGGCGATGAGAGAG 96
DB 1 ATGGGAATTAAGGCCCTGGCCAACTAATGCTGATGTGGCCCCCGACGTCCGGGAG 60
QY 97 CAGAAGTTCAGAGACTACTTCGGCCGCAAAATGCGCGTGCAGCGCAGCATATATAC 156
DB 61 AATGACATCAAGAGCTACTTGGCCGTAAGTGGCCATTGATGCTCTAATGACATTTAT 120
QY 157 CAGTTCCTGATTTGATGGAAGACAGCATGGAACCTCTCAAAATGAGCTGTGAA 216
DB 121 CAGTTCCTGATTTGCTGCTCCGAG--GTTGGGGATGTGCTGCAGATGAGGAGGTGAG 177
QY 217 GTCCTGATTTGCAAGGAATGTTCAACCGGACATAAGTATTACTGGAAGCGGGAATC 276
DB 178 ACCACGACCACTGATGAGGATGTTTACCGCACCTTGGCATGATGAGAAACGGCATC 237
QY 277 AAGCCAGTTTATGTTTGAATGGCAAGCTCTGATATGAAGAACAAGACTTGTCTAA 336
DB 238 AAGCCGCTGATGCTTGTGATGGCAAGCCGACAGCTCAAGTCAGCGCCAGCCAA 297
QY 337 AGATACCTCAAAAGAGATGATGCACCAAGATCTGACTGAGCGCATGAGGTAGAGAT 396
DB 298 CGCAGTAGCGGCGGCTGAGGCGAGAGACAGCTGCAGCGCTCAGCTGCTGGGGCC 357
QY 397 AAGATGCGATTGAAATTTGAGCAAGAGACTGTAAGGTCTACAAAGCAACAAGAA 456
DB 358 GAGCAGAGAGGTGGAATATCTACTAAGCGCTGTAAGTCTACTAAGCAGCAATGAT 417
QY 457 GATTGTAAGCGCTATTAAAGCTTATGAGGGTCTCTGTGTAAGAGCACTTCTGAAG 516
DB 418 GAGTCAAAACATCTGCTGAGGCTCATGAGCATCCCTTATCTTGTATGCCACAGTGA 477
QY 517 GAACAGAGATGTGAGGCGCTTTCATTAAGCATTAAGGTGTTGCTGTTTGAAGAT 576
DB 478 GAGGCGACGCTGCTGCGCTGTAAGGCTGCAAAATCTATGTTGGGCTTACGAGAG 537
QY 577 ATGACCTCCCTTACTTTTGGGGCTCAAGGTTCTCTGCTGATTTAATGATCAAGTTC 636
DB 538 ATGACCTCCCTTACTTTTGGGGCTCAAGGTTCTCTGCTGATTTAATGATCAAGTTC 636
QY 637 AAGAAATACCTGATGAGATTTGATGTTGCCAAGGTTTGGAGGAGCTTGAACCTAC 696
DB 598 AAAAAGCTGCAATTCAGGAATTCACCTGAGCGGATTTGCGAGGAGCTGGGCTGAAC 657
QY 697 ATGACGACGCTCATTTGATTTGATGATCCTGCTGATGCTGCTGATGATGATGATG 756
DB 658 CAGGACAGCTTTGATGATCTGCTGATGCTGCTGATGATGATGATGATGATGATG 717
QY 757 GGTATCGGGGGGCAAAAGCTCTGAATTTATTCGTAACATGGGCTCATAGAAAGCATC 816
DB 718 GGTATTTGGGCGCAAGCGGCTGTGACCTCATTCAGAAAGCAGAGACATGAGAGATG 777
QY 817 TTGAGATCTTAAATGAAGATATCAATTTCTGAGGAGCTGCGCTTACCAAGAGCT 876

```

```

DB 778 GTGCGGAGACTTGAACCCCAACAGTACCTGTGCGCAGAAAATTTGGCTCCCAAGAGAGCT 837
QY 877 CGAGCCTTTGTTCAAGAGACCTTAATG---TCACATTTGATATTTCCGTAGCTAAATGAGCT 933
DB 838 CACGAGCTCTTTGGAACCTGAGGCTGCTGAGCCCAAGAGTCTGTGAGCTGAAGTGAAGC 897
QY 934 GCACCTGATGAGAGAGGCTCTCATTAAGTTTCTGTTAAAGTAAATGTTTCAACGAAGAT 993
DB 898 GAGCCAAATGAAGAAGAGCTATCATGTTCAATGTGTGTAAGAAAGCAGTTCTCTGAGAG 957
QY 994 CGGATGACAAAGCCATAGAGACATCAATCTCCCAAGATTAATGCTGCGCAAGAGAGA 1053
DB 958 CGAATCCCGAGTGGGCTCAAGAGGCTGAGTAAGAGCCGCCAAGGACAGCACCGAGGCCGC 1017
QY 1054 CTCGAGCTCTTTTCAAGCAACTGCGCACCATCAGCACC 1094
DB 1018 CTGGATGATTTCTTCAAGGTACCGGCTCACTCTTCTTACG 1058

RESULT 7
US-08-455-968E-9
Sequence 9, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2033 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 104..1237
US-08-455-968E-9

Query Match
Best local Similarity 20.0%; Score 276.6; DB 2; Length 2033;
Pred. No. 1.7e-72;
Matches 638; Conservative 0; Mismatches 479; Indels 24; Gaps 5;

QY 37 ATGGGCAATCAAGGCTTGAAGAACTGCTGGCGGACATGCGCCCAAGGCGATGAGAGAG 96
DB 104 ATGGAATTCAGGCGCTTCCAAACTAATGCTGATGTGGCCGCCCAAGGCTCCGTAG 163
QY 97 CAGAACTTGAAGAGCTACTTGGCGGCAAAATCGCGCTGAGCGCAGCATGACATATAC 156

```


Db 164 AATGACATCAAGAGTACTTGGTGGCAAGGCCATGAGCTTCATGAGCATCTAC 223
QY 157 CAGTTCTGATTGTTATGTAAGAGAGCATGAACTCTCACAATGAAGCTGTGAA 216
Db 224 CAGTTCTGATTGTTATGTAAGAGAGCATGAACTCTCACAATGAAGCTGTGAA 280
QY 217 GTCAGTACTGATTTGCAAGAGATGTTCAACCCGAGCAATAGATTTCTGGAAGCCGGAATC 276
Db 281 ACACACAGCC--TGATGGGCAATGTTCTACCTACCATGCG--CATGGAGAAATGGCATC 334
QY 277 AACCCAGTTTATGTTTGTATGAGCAAGCTCCTGATATGAGAAACAAGAGCTGTCTAAA 336
Db 335 AACCTGTGTGATCTTTTATGAGCAAGCTCCTGATATGAGAAACAAGAGCTGTCTAAA 394
QY 337 AGATTAAGTCAAAAAGAGATGATGCAACCAAGATGATGAGAGAGAGAGATGAGAT 396
Db 395 CGCAGTGAAGAGCGCGCGCGAGGCTGAGAAAGCACTGACAGAGCTCAGCAGGCTGGATG 454
QY 397 AAGATGCGGATTAATAAATTGAGCAAGAGAGCTGTAAGAGTCAAGAGCAACACAGCA 456
Db 455 GAGAGAGAGGTGAGAAAGTTCACCAAGAGGCTGGAAGGTCAACCAACACCAATGAT 514
QY 457 GATTGTAAGCGGCTATTAAAGTATGAGAGTTCCTGTTGTAGAGAGAGCTTCTGAACA 516
Db 515 GAGTGAACCAAGCTGCTGAGAGCTCATGGGCAATCCCTTACCTTGATGACACCAAGAGCA 574
QY 517 GAAGCAAGATGTCAGACCTTTGCAATAAAGTATGAGAGTTCCTGTTGTAGAGAT 576
Db 575 GAGGCGAGCTGCTGCTGCTGCAAGAGGCTGCAAGTCTATGAGTCAAGAGAGAC 634
QY 577 ATGAGTCCCTTACTTTTGGGCTCCACAGTCTCTGCTGATTTAATGATCAAGTTC 636
Db 635 ATGAGTCCCTTACTTTTGGGCTCCACAGTCTCTGCTGATTTAATGATCAAGTTC 694
QY 637 AAGCAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
Db 695 AAGCAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 754
QY 697 ATGAGTCCCTTACTTTTGGGCTCCACAGTCTCTGCTGATTTAATGATCAAGTTC 756
Db 755 CAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 814
QY 757 GATATGAGGAGGCAAGAGCTGGAAGTATGATGATGATGATGATGATGATGATGATGAT 816
Db 815 GATATGAGGAGGCAAGAGCTGGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 874
QY 817 TTGAGATCTTTAATAAGAGATATCAATTTCTGAGAGAGTGGCTTACCAAGAGATC 876
Db 875 GTGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 934
QY 877 GAGGCTGTTGCAAGAGCTTAA--TGTCAATTTGATATTTCTGAGCTTAAATGAGCT 933
Db 935 CACAGCTCTTCTGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 994
QY 934 GCACCTGATGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993
Db 995 GAGCAATATGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1054
QY 994 CGGCTGCAAGAGCTTGAAGAGTCAATTTGCAAGATTAATTTCTGCAAGAGAGAG 1053
Db 1055 CGAATTCGAGTGGGTGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1114
QY 1054 CTGAGTCTCTTTTCAAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1113
Db 1115 CTGAGTCTCTTTTCAAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152
QY 1114 TCGGATTAAGCAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
Db 1163 AAGGAGCCAGAACCCCAAGAGGCTGCTAGAGAAAGCAAGAGTGTGAGAGAGAG 1222
QY 1174 T 1174

Db 1223 T 1223
RESULT 8
US-08-455-968E-4
Sequence 4, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1930 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-455-968E-4
Query Match 19.7%; Score 272.2; DB 2; Length 1930;
Best Local Similarity 55.1%; Pred. No. 3.4e-71;
Matches 629; Conservative 0; Mismatches 488; Indels 24; Gaps 4;
QY 37 ATGGGATCAAGGCTTTGACGAAGTCTGGCGGACAAATCGGCCGACAGCCAGCATGAAGAG 96
Db 1 ATGGGAATTCAGGGCTTGGCCAAACTAATTCGTATGTGTGGCCCCAGTGGCATCCGTGAG 60
QY 97 CAGAAATTTGAGAGTACTTGGCGGCAAAATCGCGCTGACAGCCAGCATGAGCATATAC 156
Db 61 AATGAGATCAAGAGCTACTTGTGTGTAAGTGGCATGATGCTCATGATGAGCATCTAC 120
QY 157 CAGTTCTGATTTGATGTAAGAGAGCATGAACTCTCACAATGAAGTGTGAA 216
Db 121 CAGTTCTGATTTGATGTAAGAGAGCATGAACTCTCACAATGAAGTGTGAA 177
QY 217 GTCAGTACTGATTTGCAAGAGATGTTCAACCCGAGCAATAGATTTCTGGAAGCCGGAATC 276
Db 178 ACACACAGCC--TGATGGGCAATGTTCTACCTACCATGCG--CATGGAGAAATGGCATC 231
QY 277 AAGCCAGTTTATGTTTGTATGAGCAAGCTCCTGATATGAGAAACAAGAGCTGTCTAAA 336
Db 232 AAGCTGTGTAGCTTTTATGAGCAAGCTCCTGATATGAGAAACAAGAGCTGTCTAAA 291
QY 337 AGATTAAGTCAAAAAGAGATGATGCAACCAAGATGATGAGAGAGAGATGAGAT 396
Db 292 CGCAGTGAAGAGCGCGCGAGGCTGAGAAAGCACTGACAGAGCTCAGAGAGCTGGAGT 351

OY		397	AAAGATCGAATTGAAAAATTGACCAAGGACTGTAAAGCTTACAAGGCACACACGA	456
Dd		352	CAGAGAGAGTGAGAGACTTCACCAGAAGGCTCTGTAAAGTCCAGAACACATCAT	411
OY		457	GATTGTAAAGGGCTATTTAAGACTTTATGAGGGGTTCCTGTTGTAGAGGCACTTCTGAA	516
Dd		412	GAGTGCAAAACACTCGTAGACCTCATTAGGGGATCCCTTACTCTTGAGAGCCACGAGG	471
OY		517	GAAGCAGAAATGTGACGCCCTTGTGCATAAAGATGAAGTGTGGTGGCTTCAGAMAT	576
Dd		472	GAGGCCACGCTGTGCTGCCCTCGCAAAAGCTGGCAAAAGTATATGTGCGGCCACGAG	531
OY		577	ATGGAATCCCTTACTTATTTGGGGCTCCACAGGTTCTTTCGTATTTANTGATCCAA	636
Dd		532	ATGAGACTGCGCTCACTTTTGTGGCACCCCCTGCTAATCGACACTTAATCCGAGT	591
OY		637	AAGAAAATACCTGTGATGAAATTTGATGTGGCCAAAGTTTGSAGAGGTTGAATCC	696
Dd		592	AAGAAGCTGCCCATCCCAAGAGTTCATCTGAGCGGCTCTCGAGAGAGCTGGGT	651
OY		697	ATGAGCAAGTTCATTTGATTTTGTCTATCTGTGTGTGATGTACTATTGTGATAGCA	756
Dd		652	CAGGAGAGAGTTGTGATCTGTGCATCTGCTGGTGGTAGGAGCTACTGGGAGAGCT	711
OY		757	GGTATCGGGGGGCAACAGCTCTGAACCTTATTCGCACATGGGTCATGAAAGCATC	816
Dd		712	GGCATTTGGCCCAAGCGGGCTGTGATCTTCATCCAGAAACATBAAGAGCATGAG	771
OY		817	TTTGAGAAATCTTAATAAAGACAGATATCAAATTCCTGAGAGACTTGCCCTTA	876
Dd		772	GTGAGGGGGCTGGGACCCACCAAGATGACCCCGTTCAGAAATCGCTCCACAGGA	831
OY		877	CGAGGCTGTTTCAAGGAGGCTTAA--TGTACATTTGGATTTCTGAGATTAATAG	933
Dd		832	CAGAGAGCTCTTCCGTGAGGCCAAGAGTAGTGGAGCCAGABGTCTGTGAGACTGA	891
OY		934	GCACCTGATGAGGAGGCTCTCATTAAGTTTCTGTGTAAAGATTAATGTTTTCA	993
Dd		892	GAGCCAATGAAAGAAAGTTGGTCAAAATTTATGTGTGAAAAOCAGTTTCTGAA	951
OY		994	CGGCTGACAAAGGCCATAGAGAGATCAATATGCCCCAAMAATTAATCGCCGAG	1053
Dd		952	CGAATTTGGCAGGTGGGGTCAAGCGGGCTGAGTAAGAACGCCGCAAGGACCA	1011
OY		1054	CTCAGATCTTTTTCAAAGCCAACTGCCACACATKAGCACCCGTTAAACGAGAG	1113
Dd		1012	CTCGATGATTTCTTCAAGGTGACAGGCTCACTCTCTCAGC-----TAA	1059
OY		1114	TGCGATAAACCAAGCAGGAGCTGTGCAACAAAGAAACCAAGGCTGTGGAAAG	1173
Dd		1060	AAGGAGCCAGAACCCAAAGGGGGCTGCTAAGAAAGAAAGCAAAAGCATGTGGG	1119
OY		1174	T 1174	
Dd		1120	T 1120	

RESULT 9
US-08-435-968E-6
Sequence 6, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chi-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

```

? ZIP: 4411-3834
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/455,968E
? FILING DATE: 30-MAY-1995
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Apple, Randolph T.
? REGISTRATION NUMBER: 36,429
? REFERENCE/DOCKET NUMBER: 18985-000100
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-576-0200
? TELEFAX: 415-576-0300
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1149 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? US-08-455-968E-6

```

Query Match	19.6%	Score 270.2	DB 2	Length 1149
Best Local Similarity	55.5%	Pred. No. 1e-70		
Matches 599	Conservative	0	Mismatches 433	Indels 27
				Gaps 3
QY	37	ATGCGCATCAAGGGTTTGACGAACATGCTGCGGCAATGCGCCCAAGGCGATGAGAG	96	
Db	1	ATGGGATTTAAAGGTTTGAATGCAATTAATATGGAACAATGTTCTCTGCTATGAGAA	60	
QY	97	CAGAGATTCCAGAGACTTTCGGCCCAAAATCGCCGTCGACGCGCAGATGACATATAC	156	
Db	61	AGGCATATCAAGAGCTTTTGTGSCAAAGGTTGGCCATCGATGCTCATATGTCATAT	120	
QY	157	CAGTTCCTGATTTAGTTGGAAGAGACGAGCATGSAACCTGTCACAAATGAAGCTGGGA	216	
Db	121	CAGTTTTTAATTTGCTGTGAACACGCAAGCGGTGGGACGTTGACCATGAAGCCGGTGA	180	
QY	217	GTCACATGTCATTTTGCAGAGATGTTCAACCGGCAATTAAGATTACTGGAAGCGGATC	276	
Db	181	ACAACGTCACACTGTGATGGGTATGTTTATAGACACCTGAGAAATGATTGAATACG	240	
QY	277	AAGCATTTTATGTTTTGTATGCGCAAGCCTCTGATATGAGAAACAAGACTTGTA	336	
Db	241	AAGCTTGTTATGTCTTGACGSCAACTCCAGCTTGAATATCATGATGATGACAAAG	300	
QY	337	AGATATCAAAAAAGATGATGCAACCAAGATCTGACTGAGGCATTAAGATAGGAT	396	
Db	301	CGGCTTCAAGAAAGGTTGGAACGAAACGAAACGAGGCAACACAGATTTGAA	360	
QY	397	AAATATCGATTTAAAAAATTGAGCAGAGAGACTGTAAAGGTGCACAGGCAACACAGAA	456	
Db	361	AAATGTAAGCAAGAAGA-----ACATTGTTGAAGGTCCTCAAAAGACATTAATGAA	411	
QY	457	GATTGTAAACGGCTATTAGACTTATGSSGGTTCCTGTTTGAAGGCACTTGAAGCA	516	
Db	412	GAAGCCCAAAATTAATCTAGAGACTTAATGGGAATCCCATATTAATAGGCCCAACGAAAGCT	471	
QY	517	GAACGCAATGTCCAGGCCCTTTGCATAAACGATAAGGTGTTGCGTGTTCCTCAAGAT	576	
Db	472	GAGCGCTAAATGTCTGTGATGGCAAAAGAGGGAAGGATGTGCGCGAGCAAGTAAGAT	531	
QY	577	ATGAGTCCTCTTACTTTTGGGGCTCCACGGTTCCTGTCATTTAAAGATATCCAAAGTTC	636	
Db	532	ATGACACACTCTGTTATAGAACACCCTCTTCTTGTTGAACACTTTGCTTTTCAGAGGC	591	
QY	637	AAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGAGAGACTTGAATCACC	696	
Db	592	AAGGAAGGACCGATTACAGCAATAGATGATGATTAATGTTTGGAGAGGACCTGCATTCACA	651	


```

; APPLICANT: Kosugi, Yoshitsugu
; APPLICANT: Matsui, Eriko
; APPLICANT: Kawasaki, Satoko
; TITLE OF INVENTION: THERMOSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: HYPERTHERMOPHILE BACTERIUM BELONGING TO THE GENUS
; TITLE OF INVENTION: PYROCOCUS
; FILE REFERENCE: 081356-0121
; CURRENT APPLICATION NUMBER: US/09/146,319
; EARLIER APPLICATION NUMBER: 1998-09-03
; EARLIER FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Pyrococcus horikoshii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
; US-09-146-319-1

```

```

Query Match          9.8%; Score 135.6; DB 4; Length 1032;
Best Local Similarity 52.1%; Pred. No. 1.2e-30;
Matches 303; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

QY 86 CGATGAGGAGCAGAGATTGAGAGCTACTCGGCCGCAAAATCGCGTCGAGCCGACGA 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29 CGAGGAGGAGATAGATCTTTGAAATCTGTATGGAAAGAGATAGCGATGATGCCCTAA 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 146 TGAGCATATACCAAGTTCCTGATTAGTTGGAAGAGACAGGAGATGGAACCTGCACAATG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 AGCCCATATACAGTTTATATCAAGATAGACAGAGGAGATGGAACCACTTATGAGACT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 AGCTGGTGAAGTCACTACTCTATTGCAAGATGTTAACCGGACATTAAGTTACTGG 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 149 CTAAAGGTAGGATACCTCTCTATTAAGTGGCTCTTTTATAGAACGTAATATCTAATGG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 AAGCGGAGATCAAGCAGATTATGTTTGTATGAGCAGCCTCTGATGTAAGTAAGCAAG 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 AAGCCGGATTATAGCGCGCTACGCTTTGATGGAAGACCTCCGGAATTCAGAAAGGAAG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 AGCTTGTAAAAGATATCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGCAGTAG 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 AGCTCGAAAAAAGAGAGAGAGCTAGAGAAAGAGCAAACTAAATGGAAGAAAGCTCTAG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 AGGTAGAGATTAAGATGCGATTGAAAAATTGAGCAGAGAGACTGTAAAGTCCAGAGGC 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 CCAAGGAAACCTGAGAGAGCTAGAGAAATACGTCGAAGGGCAACTAAGTTAATGAAA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 AACACAGCAGAGATTGTAAACGGCTATTAGACTTATGGGGTTCCTGTTAGAGGCAC 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 TGCTAATCGAAGATGCAAGAAAGCTTTTGCAACTAATGGGAATACCATTAATTCAGCCTC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 CTTCGAGCAGACAGAGATGTCAGCCCTTTCATTAAGATTAAGTATGCTGCTGTTG 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 CAAGTGAAGAGAGAGAGAGAGAGAGATGATGCAAGTAAGGGATGTCTACGCGCTAG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 CTTCAGAGATATGATGCTCTTACTTTTGGGCTCCACGGTTCCTTCGTCATTTAATGG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CGAGTCAAGATTATGATTAATCTACTACTCTTTGGTGCTCCAAAGGTGATTAAGAAATCTGACAA 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626 ATCCAAAGTTCAGAAAGAAATACCTGTGATGGAATTTGATGTTG 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 TTACGGGAAAAAGAAAGATGCTGGGAAAAAGATGTTTACGTTG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-09-175-973-1
; Sequence 1, Application US/09175973A
; Patent No. 6255081
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO

```

```

; APPLICANT: ISHIKAWA, Kazuhiko
; APPLICANT: KOSUGI, Yoshitsugu
; APPLICANT: MATSUI, Eriko
; APPLICANT: KAWASAKI, Satoko
; TITLE OF INVENTION: THERMOSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: HYPERTHERMOPHILE BACTERIUM BELONGING TO THE GENUS
; TITLE OF INVENTION: PYROCOCUS
; FILE REFERENCE: 081356/0126
; CURRENT APPLICATION NUMBER: US/09/175,973A
; EARLIER FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Pyrococcus horikoshii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
; US-09-175-973-1

```

```

Query Match          9.8%; Score 135.6; DB 4; Length 1032;
Best Local Similarity 52.1%; Pred. No. 1.2e-30;
Matches 303; Conservative 0; Mismatches 279; Indels 0; Gaps 0;

QY 86 CGATGAGGAGCAGAGATTGAGAGCTACTTCGCCGCAAAATCGCGTCGAGCCGACGA 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29 CGAGGAGGAGATAGATCTTTGAAATCTGTATGGAAAGAGATAGCGATGATGCCCTAA 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 146 TGAGCATATACCAAGTTCCTGATTAGTTGGAAGAGACAGGAGATGGAACCTGCACAATG 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 AGCCCATATACAGTTTATATCAAGATAGACAGAGGAGATGGAACCACTTATGAGACT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 AGCTGGTGAAGTCACTACTCTATTGCAAGATGTTCAACCGGACATTAAGTTACTGG 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 149 CTAAAGGTAGGATACCTCTCTATTAAGTGGCTCTTTTATAGAACGTAATATCTAATGG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 AAGCGGAGATCAAGCAGATTATGTTTGTATGAGCAGCCTCTGATGTAAGTAAGCAAG 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 AAGCCGGATTATAGCGCGCTACGCTTTGATGGAAGACCTCCGGAATTCAGAAAGGAAG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 AGCTTGTAAAAGATATCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGCAGTAG 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 AGCTCGAAAAAAGAGAGAGAGCTAGAGAAAGAGCAAACTAAATGGAAGAAAGCTCTAG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 AGGTAGAGATTAAGATGCGATTGAAAAATTGAGCAGAGAGACTGTAAAGTCCAGAGGC 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 CCAAGGAAACCTGAGAGAGCTAGAGAAATACGTCGAAGGGCAACTAAGTTAATGAAA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 AACACAGCAGAGATTGTAAACGGCTATTAGACTTATGGGGTTCCTGTTAGAGGCAC 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 TGCTAATCGAAGATGCAAGAAAGCTTTTGCAACTAATGGGAATACCATTAATTCAGCCTC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506 CTTCGAGCAGACAGAGATGTCAGCCCTTTCATTAAGATTAAGTATGCTGCTGTTG 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 CAAGTGAAGAGAGAGAGAGAGAGATGATGCAAGTAAGGGATGTCTACGCGCTAG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566 CTTCAGAGATATGATGCTCTTACTTTTGGGCTCCACGGTTCCTTCGTCATTTAATGG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CGAGTCAAGATTATGATTAATCTACTACTCTTTGGTGCTCCAAAGGTGATTAAGAAATCTGACAA 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626 ATCCAAAGTTCAGAAAGAAATACCTGTGATGGAATTTGATGTTG 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 TTACGGGAAAAAGAAAGATGCTGGGAAAAAGATGTTTACGTTG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 5, 2002, 11:03:04
Job time : 44.2033 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:28:52 : Search time 1181.61 seconds
(without alignments)
15774.429 Million cell updates/sec

Title: US-09-805-311-5
Perfect score: 1381
Sequence: 1 cgaccacgcgcgcgcac.....ltgaaaaaaaaaaaaaa 1381

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estda:
2: em_esthum:
3: em_estin:
4: em_estnu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_hlc:
9: gb_estl:
10: gb_estl2:
11: gb_hlc:
12: gb_gss:
13: em_gss_hum:
14: em_gss_iny:
15: em_gss_pln:
16: em_gss_vtc:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	770.2	55.8	901	10	BG837708 Zm10_01f0
2	523.8	37.9	550	10	BE639422 946033A02
3	511.2	37.0	554	10	BE639421 946033A02
4	509	36.9	586	9	AI881599 606068G09
5	501	36.3	532	9	AM000375 614014D03
6	497.8	36.0	553	9	AM562789 660065H06
7	465.2	33.7	470	9	AI861468 614014D03
8	463.4	33.6	474	9	AM539173 660065H06
9	458.2	33.2	648	9	AV913663 946012C08
10	448	32.4	467	9	BE186786 946012C08
11	420.4	30.4	733	9	AM774700 EST333851
12	411.6	29.8	539	10	BG263333 WHR2339_F
13	409.8	29.7	475	9	AM562517 660065H06
14	402.2	29.1	414	9	AM288784 707010C02
15	397	28.7	456	9	AI065689 ag91f12.x
16	393	28.5	553	9	AI834484 606068G09
17	378.8	27.4	632	9	AM585913 EST317536

18	372.6	27.0	609	10	BG096545
19	370.4	26.8	734	10 <td>BE416584</td>	BE416584
20	369.6	26.8	820	10 <td>BG414505 HVSMK000</td>	BG414505 HVSMK000
21	364.6	26.4	645	9	AM691681 NF043B09S
22	348	25.2	460	10 <td>BE417817</td>	BE417817
23	347.2	25.1	683	9	AV918966
24	346.8	25.1	454	9	AV915644
25	325.4	23.6	539	10 <td>BP098320</td>	BP098320
26	321.8	23.3	363	9	AM562518
27	317.2	23.0	470	9	AM288831
28	314.8	22.8	376	9	AI065546
29	314.8	22.8	579	9	AV833393
30	314.4	22.8	592	9	AI780966
31	308	22.3	526	10 <td>BI419720</td>	BI419720
32	300.6	21.8	695	9	BE052579
33	273.8	19.8	942	9	AL560007
34	273	19.8	549	9	AM696665
35	265.6	19.2	903	9	AL531350
36	263.8	19.1	851	9	AL519300
37	248.6	18.0	689	10 <td>BE311755</td>	BE311755
38	245.4	17.8	769	10 <td>BG574950</td>	BG574950
39	244.4	17.7	660	10 <td>BJ011893</td>	BJ011893
40	244.2	17.7	806	9	AL560395
41	237.6	17.2	762	10 <td>BG337603</td>	BG337603
42	236	17.1	866	10 <td>BI827898</td>	BI827898
43	234.4	16.9	785	10 <td>BG756459</td>	BG756459
44	233.6	16.9	704	9	AM783795
45	233.4	16.9	963	12 <td>CNS06WXR</td>	CNS06WXR

ALIGNMENTS

RESULT 1
LOCUS BG837708 901 bp mRNA linear EST 25-MAY-2001
DEFINITION Zm10_01f08.A Zm10-AAFC_ECORC_Fusarium_graminearum_corn_silk Zea
mays cDNA clone Zm10_01f08, mRNA sequence.
ACCESSION BG837708.1 GI:14204031
VERSION
KEYWORDS
SOURCE EST
ORGANISM Zea mays.
Zea mays.
Zea mays.

REFERENCE
AUTHORS Harris,L.J., Balcerzak,M., Allard,S., Sapano,A., Couroux,P., De Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A., Sprott I (bases 1 to 901)
Clade: Panicoideae; Andropogonaceae; Zea.

TITLE
JOURNAL
COMMENT
Expressed Sequence Tags from Maize Silk Six Hours After Silk Channel Inoculation with Fusarium graminearum
Unpublished (2001)
Contact: Harris, Linda J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1314
Fax: (613) 759-6566
Email: harris.lj@em.agr.ca.

FEATURES
Source
Location/Qualifiers
1..901
/organism="Zea mays"
/cultivar="CO388"
/db_xref="taxon:4577"
/clone="Zm10_01f08"
/clone_lib="Zm10-AAFC_ECORC_Fusarium_graminearum_corn_silk"

"/tissue_type="Silk"
/dev_stage="4-5 days post-silk emergence"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI; Site 2: XhoI; Field-grown corn was silk channel-inoculated

In the morning (~10 am) with 1 ml of a macroconidial suspension (500,000 spores/ml) of *Fusarium graminearum* and silk channels were collected and immediately frozen in liquid nitrogen 6 hours later. RNA was extracted from silk tissue between 1 cm below and above the inoculation point in the silk channel, RNA from five silk channels was pooled."

BASE COUNT 267 a 208 g 249 t 3 others
ORIGIN 174 c

Query Match 55.8%; Score 770.2; DB 10; Length 901;
Best Local Similarity 98.3%; Pred. No. 1,2e-153;
Matches 797; Conservative 2; Mismatches 10; Indels 2; Gaps 2;

QY 557 TCCTGTTGCTTCAGAGATATGAGCTCCCTTACTTTGGGGCTCCAGGTTCTCTGTC 616
DB 1 TCCTGTTGCTTCAGAGATATGAGCTCCCTTACTTTGGGGCTCCAGGTTCTCTGTC 60
QY 617 ATTAAATGATCCAAAGTCCAAAGAAATACCTGTATGAAATTGATGTGCCAAGGTTT 676
DB 61 ATTAAATGATCCAAAGTCCAAAGAAATACCTGTATGAAATTGATGTGCCAAGGTTT 120
QY 677 TGGAGAGCTTGAACCTCACCATGACACATATTGATTGTGTCATCTCTGTGTGATGTG 736
DB 121 TGGAGAGCTTGAACCTCACCATGACACATATTGATTGTGTCATCTCTGTGTGATGTG 180
QY 737 ACTATGTGATAGCATCAAAAGTATCGGGGGCAACAGCTCTGAAACTTATTCGTACAC 796
DB 181 ACTATGTGATAGCATCAAAAGTATCGGGGGCAACAGCTCTGAAACTTATTCGTACAC 240
QY 797 ATGGTCCATAGAAAGATCTTGAGATCTTAAATTAAGACAGATCAAAATTCMTGAG 855
DB 241 ATGGTCCATAGAAAGATCTTGAGATCTTAAATTAAGACAGATCAAAATTCMTGAG 300
QY 856 GACTGGCTTACC - AAGAACTCTGACGCTGTGTCAAGAGCGCTAAATGTACATTTGATAT 914
DB 301 GACTGGCTTACC - AAGAACTCTGACGCTGTGTGTCAAGAGCGCTAAATGTACATTTGATAT 360
QY 915 TCCGAGCTAAATGAGCTGACCTGATGAGAGGGTCCATTAATCTCCGTGTAAGA 974
DB 361 TCCGAGCTAAATGAGCTGACCTGATGAGAGGGTCCATTAATCTCCGTGTAAGA 420
QY 975 TAATGTTTCAACGAAGATGGGTGACAAAGCCATAGAGATCAAAATCTGCCAAGAA 1034
DB 421 TAATGTTTCAATGAAGATGGGTGACAAAGCCATAGAGATCAAAATCTGCCAAGAA 480
QY 1035 TAAATGTCGCAAGGAAGACTCGAGTCTTTTCAAGCAACTGTCACACATCAGACGC 1094
DB 481 TAAATGTCGCAAGGAAGACTCGAGTCTTTTCAAGCAACTGTCACACATCAGACGC 540
QY 1095 GCTAAACGAGAGAGACTTGGATTAACAAGCAAGGAGCTGCGAACAAGAAACAAA 1154
DB 541 GCTAAACGAGAGAGACTTGGATTAACAAGCAAGGAGCTGCGAACAAGAAACAAA 600
QY 1155 GCGTGTGGAAGAAAGAAATATCTTGATGCTGTGATGACAACTACAGACTACGAAAGCA 1214
DB 601 GCGTGTGGAAGAAAGAAATATCTTGATGCTGTGATGACAACTACAGACTACGAAAGCA 660
QY 1215 GCGGTGCGTATCACTTGGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGG 1274
DB 661 GCGGTGCGTATCACTTGGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGG 720
QY 1275 TAAAGTTGCGCATGTTTAAAGCTGGGGTAAGTTAGTTGTTTGAAGAAATTTGGGTA 1334
DB 721 TGAAGTTTGGCCATGTTTAAAGCTGGGGTAAGTTAGTTGTTTGAAGAAATTTGGGTA 780
QY 1335 CCAAGTAAACAAACTTATCGCTGTTTTTTGA 1365
DB 781 CCAAGTAAACAAACTTATCGCTGTTTTTTGA 811

RESULT 2
BE639422

LOCUS BE639422 550 bp mRNA linear EST 30-AUG-2000
DEFINITION 946033A02.y2 946 - tassels primordium prepared by Schmidt Lab Zea
ACCESSION mays CDNA, mRNA sequence.
VERSION BE639422
KEYWORDS BE639422.1 GI:9952839
SOURCE EST.
ORGANISM Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 550)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.
Location/Qualifiers
FEATURES
source
1..550
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:457"
/clone_lib="946 - tassels primordium prepared by Schmidt Lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybriZAP; site_1: EcoRI; site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."
BASE COUNT 177 a 116 c 149 g 108 t
ORIGIN
Query Match 37.9%; Score 523.8; DB 10; Length 550;
Best Local Similarity 98.5%; Pred. No. 2.9e-101;
Matches 539; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1 CGACCCACGGCGTCGCGCACAGCGCCGCGACAGAGATGGCGATCAAGGTTTGACGAA 60
DB 4 CGCGCCCGCCACCGCCGACAGCGCCGCGACAGAGATGGCGATCAAGGTTTGACGAA 63
QY 61 CTGCTGGCGGACAAATGCGCCCAAGGCGATGAAGAGCAGAAATTCGAGAGCTTCTGGC 120
DB 64 CTGCTGGCGGACAAATGCGCCCAAGGCGATGAAGAGCAGAAATTCGAGAGCTTCTGGC 123
QY 121 CGCAAAATCGCGCTGCGAGCGCCAGCATGAGATATACCATTCGATGTGTGGAAG 180
DB 124 CGCAAAATCGCGCTGCGAGCGCCAGCATGAGATATACCATTCGATGTGTGGAAG 183
QY 181 ACAGG - CATGGAACCTCTCAAAATGAAGCTGTGAAGTCACTAGTCATTTGCAAGAAAT 239
DB 184 ACTGTCATGGAACCTCTCAAAATGAAGCTGTGAAGTCACTAGTCATTTGCAAGAAAT 243
QY 240 GTTCAACCGGACATTAAGTATCTGAGAGGGGAAATCAAGCCAGTTTATGTTTGAATG 299
DB 244 GTTCAACCGGACATTAAGTATCTGAGAGGGGAAATCAAGCCAGTTTATGTTTGAATG 303
QY 300 CAAGCTTCCTGATATGAAGAAACAGAGCTTGTCTAAAGATCTCAAAAGAGATGATGC 359
DB 304 CAAGCTTCCTGATATGAAGAAACAGAGCTTGTCTAAAGATCTCAAAAGAGATGATGC 363
QY 360 AACCAAGATCTGACGTGAGGCTAGAGGTAGAGTAAGATGAGTGAATGAATGAG 419

Db 364 AACCAAGATCTGACTGAGCAGTAGAGATTAAGATGCGATTGAAAAATTGAG 423
 QY 420 CAAGAGACTGTAAGGTCACACAGCAACAGAGATTGTAAGCGCTATTAGACT 479
 Db 424 CAAGAGACTGTAAGGTCACACAGCAACAGAGATTGTAAGCGCTATTAGACT 483
 QY 480 TATGGGGTTCCTGTTGTAGAGCAGCTTCCTGAGCAGAGAGATGAGCCCTTGG 539
 Db 484 TATGGGGTTCCTGTTGTAGAGCAGCTTCCTGAGCAGAGAGATGAGCCCTTGG 543
 QY 540 CATAAC 546
 Db 544 CATAAC 550

RESULT 3
 BE639421/c 554 bp mRNA linear EST 30-AUG-2000
 LOCUS 946033A02.x2 946 - tassal primordium prepared by Schmidt lab Zea
 DEFINITION

ACCESION BE639421
 VERSION BE639421.1 GI:9952838
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 554)
 Walbot, V.

AUTHORS Zea ESTs from various cDNA libraries sequenced at Stanford
 TITLE University
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 946033 row: A column: 02.

FEATURES
 source
 1..554
 /organism="Zea mays"
 /cultivar="OH43"
 /db_xref="taxon:4577"
 /clone_lib="946 - tassal primordium prepared by Schmidt
 lab"
 /tissue_type="tassels"
 /dev_stage="just after the transition from vegetative to
 inflorescence development"
 /lab_host="XLOLR"
 /note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
 Site_2: XhoI; George Chuck dissected immature tassels
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA
 library in HybriZAP. Sample insert size range was 350 bp
 to 3 kb with a 1 kb average."

BASE COUNT 145 a 131 c 110 g 168 t
 ORIGIN

Query Match 37.0%; Score 511.2; DB 10; Length 554;
 Best Local Similarity 99.4%; Pred. No. 1.4e-98;
 Matches 513; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 848 TTCTGAGCAGCGCTTCGAAGAAGCTGAGCGCTGTTCAAGAGCCTAATGTCACAT 907
 Db 554 TTCTGAGCAGCGCTTCGAAGAAGCTGAGCGCTGTTCAAGAGCCTAATGTCACAT 495
 QY 908 TGAATATCTCTGAGTAAATGAGCTGACCTGATGAGAGGGTCTCATAACTTCTCG 967
 Db 494 TGAATATCTCTGAGTAAATGAGCTGACCTGATGAGAGGGTCTCATAACTTCTCG 435
 QY 968 TAAAGATATGTTTCAAGGAAGATCGGGTGAAGAGCCATTAAGAAATCAATCTG 1027

Db 434 TAAAGATATGTTTCAAGGAAGATCGGGTGAAGAGCCATTAAGAAATCAATCTG 375
 QY 1028 CCAAGATATATCTGTCGAGAGAACTCGAGCTCTTTTCAACCAACTGCCACCAT 1087
 Db 374 CCAAGATATATCTGTCGAGAGAACTCGAGCTCTTTTCAACCAACTGCCACCAT 315
 QY 1088 CAGCAGCGCTTAAAGGGAAGAGAGACTTCGATTAACAAGCAAGCAGCTGCGCAACAGA 1147
 Db 314 CAGCAGCGCTTAAAGGGAAGAGAGACTTCGATTAACAAGCAAGCAGCTGCGCAACAGA 255
 QY 1148 AAACCAAGCTGCTGGAAGAAGAATATCTTGATGCTGATGATGACACTACACTAC 1207
 Db 254 AAACCAAGCTGCTGGAAGAAGAATATCTTGATGCTGATGATGACACTACACTAC 195
 QY 1208 GAAGAGCGGTGGCGTATCATCTCGCTTATATTTAACTCCCTGTTTAACTCAGA 1267
 Db 194 GAAGAGCGGTGGCGTATCATCTCGCTTATATTTAACTCCCTGTTTAACTCAGA 135
 QY 1268 GCTTTGTTAAAGTTTCGCCATGTTTCAAGCTGGGCTAAGTTAGTTGTTTGAAGAT 1327
 Db 134 GCTTTGTTAAAGTTTCGCCATGTTTCAAGCTGGGCTAAGTTAGTTGTTTGAAGAT 75
 QY 1328 TGCTGACCAAGTAAACAACTTATCGCTGTTTTT 1363
 Db 74 TGCTGACCAAGTAAACAACTTATCGCTGTTTTT 39

RESULT 4
 A1881599 586 bp mRNA linear EST 02-FEB-2000
 LOCUS 606068G09.y1 606 - Ear tissue cDNA library from Schmidt lab Zea
 DEFINITION

ACCESION A1881599
 VERSION A1881599.1 GI:5566733
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 586)
 Walbot, V.

AUTHORS Zea ESTs from various cDNA libraries sequenced at Stanford
 TITLE University
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 606068 row: G column: 09.

FEATURES
 source
 1..586
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="606 - Ear tissue cDNA library from Schmidt
 lab"
 /tissue_type="mixed"
 /dev_stage="ear length from 0.5 cm - 2.0 cm"
 /lab_host="XLOLR (Stratagene)"
 /note="Organ: immature ear; Vector: pBR-CMV; Site_1: EcoRI
 ; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
 lab"

BASE COUNT 186 a 132 c 159 g 109 t
 ORIGIN

Query Match 36.9%; Score 509; DB 9; Length 586;
 Best Local Similarity 98.1%; Pred. No. 4.1e-98;
 Matches 515; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY      1  CGACCCAGCGGTCGGGCGACAGCCGCCGACAGCATGAGGCTTTCAGCAAA  60
Db      62  CGCCCGCCCGCCACCGCCGACAGCCGCCGACAGCATGAGGCTTTCAGCAAA  121
QY      61  CTGTGGCGGACAAATGCGCCCAAGGCGATGAAGAGCAGAAAGTTGAGAGCTTCCGCC  120
Db      122  CTGTGGCGGACAAATGCGCCCAAGGCGATGAAGAGCAGAAAGTTGAGAGCTTCCGCC  181
QY      121  CGCAAAATTCGCCGTGAGAGCCCGCATGAGCATATACAGTTCCTGATGTGTGTTGGAAG  180
Db      182  CGCAAAATTCGCCGTGAGAGCCCGCATGAGCATATACAGTTCCTGATGTGTGTTGGAAG  241
QY      181  ACAGGATGGAATCTCTCAAAATGAAGTGTGTAAGTCTAGTCAATTTGCAAGAAATG  240
Db      242  ACAGGATGGAATCTCTCAAAATGAAGTGTGTAAGTCTAGTCAATTTGCAAGAAATG  301
QY      241  TTCACCGGACAAATGAATTAATGATGAGAGGGAATCAACCCAGTTTATGTTTATGATGG  300
Db      302  TTCACCGGACAAATGAATTAATGATGAGAGGGAATCAACCCAGTTTATGTTTATGATGG  361
QY      301  AAGCCCTCGATATGAAGAACAAGAGCTTGTCTAAAGATCTCAAAAAAGAGATGATCA  360
Db      362  AAGCCCTCGATATGAAGAACAAGAGCTTGTCTAAAGATCTCAAAAAAGAGATGATCA  421
QY      361  ACCAAAGATCTGACTAGGAGCAGTAGAGTAGAGATAAAGATCGATTGAAAATTTGAGC  420
Db      422  ACCAAAGATCTGACTAGGAGCAGTAGAGTAGAGATAAAGATCGATTGAAAATTTGAGC  481
QY      421  AAGAGACTGTAAAGGTCAACAGGACAGCAACAGAAATTTGAACGGCTATTAGACTT  480
Db      482  AAGAGACTGTAAAGGTCAACAGGACAGCAACAGAAATTTGAACGGCTATTAGACTT  541
QY      481  ATGGGGGTTCTCTGTTGTAGAGCACCCTTCTGAAGCAGAAACAGAA  525
Db      542  ATGGGGGTTCTCTGTTGTAGAGCACCCTTCTGAAGCAGAAACAGAA  586

RESULT 5
AM000375      532 bp      mRNA      linear      EST 08-SEP-1999
LOCUS      614014D03.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION      mRNA sequence.
ACCESSION      AM000375
VERSION      AM000375.1 GI:5847296
KEYWORDS      EST.
SOURCE      Zea mays.
ORGANISM      Zea mays.
REFERENCE      Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE      clade; Panicoidae; Andropogoneae; Zea.
JOURNAL      1 (bases 1 to 532)
COMMENT      Walbot, V.
      Maize ESTs from various cDNA libraries sequenced at Stanford
      University
      Unpublished (1999)
      Contact: Walbot V
      Department of Biological Sciences
      Stanford University
      855 California Ave, Palo Alto, CA 94304, USA
      Tel: 650 723 2227
      Fax: 650 725 8221
      Email: walbot@stanford.edu
      Plate: 614014 row: D column: 03.
FEATURES
Source
      Location/Qualifiers
      1..532
      /organism="Zea mays"
      /cultivar="W23"
      /db_xref="taxon:4577"
      /clone_lib="614 - root cDNA library from Walbot Lab"
      /tissue_type="root"
      /dev_stage="3-4 days old"
      /lab_host="X10LR"

```

```

BASE COUNT      173 a      103 c      129 g      127 t
ORIGIN
Query Match      36.3%; Score 501; DB 9; Length 532;
Best Local Similarity 98.7%; Pred No. 2; Le-96;
Matches 526; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
/Note="Organ: root; Vector: pBluescriptII SK+; Site-1:
ECORI; Site-2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
QY      618  TTTAATGATCCAAAGTCCCAAGAAATTAAGTATGATGAAATTTGATGTTCCAGAGTTT  677
Db      1  TTTAATGATCCAAAGTTCCAAGAAATTAAGTATGATGAAATTTGATGTTCCAGAGTTT  60
QY      678  GGAGAGCTTGAAGTCAACCATGAGCCAGTCAATGTTGATGATGATGATGATGATGATG  737
Db      61  GGAGAGCTTGAAGTCAACCATGAGCCAGTCAATGTTGATGATGATGATGATGATGATG  120
QY      738  CTATTGTGATAGCATCAAGGTATCGGGGGCAAGACGCTGTGAAACTTATTCGTCAACA  797
Db      121  CTATTGTGATAGCATCAAGGTATCGGGGGCAAGACGCTGTGAAACTTATTCGTCAACA  180
QY      798  TGGGTCCATAGAAAGCATCTTGGAGAACTTTAATAAGACA -GATATCAAAATTCGAGG  856
Db      181  TGGGTCCATAGAAAGCATCTTGGAGAACTTTAATAAGACGATATCAAAATTCGAGG  240
QY      857  ACTGGCCTTACCAAGAGCTCGACGCTTGTCAAGAGACCTTAATGTCATTGGATATTC  916
Db      241  ACTGGCCTTACCAAGAGCTCGACGCTTGTCAAGAGACCTTAATGTCATTGGATATTC  299
QY      917  CTGAGCTTAAATGAGCTGCACCTGATGAGAGGCTCTCATTAAGTTTCTGTGTAAGATA  976
Db      300  CTGAGCTTAAATGAGCTGCACCTGATGAGAGGCTCTCATTAAGTTTCTGTGTAAGATA  359
QY      977  ATGTTTCAAGAGATCGGTTGACAAAGGCCATGAGAAATCAATCTGCCAAGATA  1036
Db      360  ATGTTTCAAGAGATCGGTTGACAAAGGCCATGAGAAATCAATCTGCCAAGATA  419
QY      1037  AATCGTCGACGAGAAAGACTCGATCTTTTCAAGCCAACTGCCACCAATCAGCACCCG  1096
Db      420  AATCGTCGACGAGAAAGACTCGATCTTTTCAAGCCAACTGCCACCAATCAGCACCCG  479
QY      1097  TAAACGAGAGAGACTTCGATTAACCAAGCAGGCTGGGACAAAGAA  1149
Db      480  TAAACGAGAGAGACTTCGATTAACCAAGCAGGCTGGGACAAAGAA  532

RESULT 6
AM562789      553 bp      mRNA      linear      EST 10-MAR-2000
LOCUS      660065H06.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION      mRNA sequence.
ACCESSION      AM562789
VERSION      AM562789.1 GI:7216667
KEYWORDS      EST.
SOURCE      Zea mays.
ORGANISM      Zea mays.
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE      clade; Panicoidae; Andropogoneae; Zea.
JOURNAL      1 (bases 1 to 553)
COMMENT      Walbot, V.
      Maize ESTs from various cDNA libraries sequenced at Stanford
      University
      Unpublished (1999)
      Contact: Walbot V
      Department of Biological Sciences
      Stanford University
      855 California Ave, Palo Alto, CA 94304, USA
      Tel: 650 723 2227
      Fax: 650 725 8221
      Email: walbot@stanford.edu
      Plate: 660065 row: H column: 06.

```


Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade: Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 474)
 Wabbot.V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 Contact: Wabbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 723 8221
 Email: wabbot@stanford.edu
 Plate: 660065 row: H column: 06.
 Location/Qualifiers
 1. 474
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /issue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="X10LR"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."
 BASE COUNT 113 a 114 c 92 g 154 t 1 others
 ORIGIN
 Query Match 33.6%; Score 463.4; DB 9; Length 474;
 Best Local Similarity 98.5%; Pred. No. 2e-88; Indels 0; Gaps 0;
 Matches 467; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 872 AAGTCGACGCTGTTCAAGAGCCCTPAATGTCATATTCCTGAGCTAAATGCA 931
 Db 474 AAGCTCAGCCCTGTTCAAGAGCCCTPAATGTCATATTCCTGAGCTAAATGCA 415
 Oy 932 CTGCACCTGATGAGAGAGGCTCTCATAGTTTCTGTTAAAGATATGTTTCAAGCA 991
 Db 414 CTGCACCTGATGAGAGAGGCTCTCATAGTTTCTGTTAAAGATATGTTTCAATGAG 355
 Oy 992 ATCGGGTGAACAAGGCCATGAGAGATCAATCTGCCAAGATTAATCTGCCAAGCA 1051
 Db 354 ATCGGGTGAACAAGGCCATGAGAGATCAATCTGCCAAGATTAATCTGCCAAGCA 295
 Oy 1052 GACTCGAGTCCTTTTCAAGCCAACTGCCACACATCAGACCGCTAAAGCGAAGAGA 1111
 Db 294 GACTCGAGTCCTTTTCAAGCCAACTGCCACACATCAGACCGCTAAAGCGAAGAGA 235
 Oy 1112 CTTCGGATTAACAAGCAAGGCAAGCTGCGAACAAGAAAACAAAGGCTGTGGAAGAAGA 1171
 Db 234 CTTCGGATTAACAAGCAAGGCAAGCTGCGAACAAGAAAACAAAGGCTGTGGAAGAAGA 175
 Oy 1172 AATAATCTGGATGCTGATGTACACTACACTACAGTACAGCAAGCGGCTGCTGATCACT 1231
 Db 174 AATAATCTGGATGCTGATGTACACTACACTACAGTACAGCAAGCGGCTGCTGATCACT 115
 Oy 1232 TCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTGGTAAAGTCCGCCATGT 1291
 Db 114 TCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTGGTAAAGTCCGCCATGT 55
 Oy 1292 TTCAAGCTGGGGTAAAGTATGTTGTTGAAGATTTGGTATCCAGTAACAA 1345
 Db 54 TTCAAGCTGGGGTAAAGTATGTTGTTGAAGATTTGGTATCCAGTAACAA 1
 RESULT 9
 AV913663 648 bp mRNA linear EST 18-JAN-2002
 LOCUS AV913663 K. Sato unpublished cDNA library, cv. Haruna Nijo
 DEFINITION germination shoots Hordeum vulgare subsp. vulgare cDNA clone

bag922n17 5', mRNA sequence.
 AV913663
 AV913663.1 GI:18209440
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Hordeum vulgare subsp. vulgare.
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 648)
 Sato, K., Saito, D. and Takeda, K.
 Barley EST sequencing project in NIG and Okayama Univ
 Unpublished (2002)
 Contact: Tadasi Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 Location/Qualifiers
 1. 648
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Haruna Nijo"
 /db_xref="taxon:112509"
 /clone_lib="bag922n17"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo germination shoots"
 /issue_type="shoots"
 /dev_stage="germination"
 BASE COUNT 174 a 158 c 185 g 131 t
 ORIGIN
 Query Match 33.2%; Score 458.2; DB 9; Length 648;
 Best Local Similarity 86.0%; Pred. No. 2.6e-87;
 Matches 508; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 Oy 13 CCGGCCACAGCCGCCGACAGAGATGGGCTTACGAAACTGTGCGGAC 72
 Db 58 CCGGTCGCCGCCGCGCGCGGATGGAGTCAAGAGGTTTACGAACCTCTGCGGAC 117
 Oy 73 AATGCCGCCAGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132
 Db 118 AAGCGGCCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 177
 Oy 193 GTGAGGCCAGATGAGATATACAGTTCTGATTTGATTTGGAAGAGAGAGATGGA 192
 Db 178 GTGAGGCCAGATGAGATATACAGTTCTGATTTGATTTGGAAGAGAGAGATGGA 237
 Oy 238 ACCCTTACAAAGCAAGCCGATGATGACAGATGATGCAAGGATGTCAGCGGACA 252
 Db 238 ACCCTTACAAAGCAAGCCGATGATGACAGATGATGCAAGGATGTCAGCGGACA 297
 Oy 297 ATAGATTACTAGAACCGGGAATCAAGCCAGTTATGTTTGTGATGAGAGAGAGATGAT 312
 Db 298 ATAGATTACTAGAACCGGGAATTAACCAAGTATGTTTGTGATGAGAGAGAGATGAT 357
 Oy 313 ATGAGAAAGAGAGCTTGTCTAAAGATATCTCAAAAAGAGATGATGCAACCAAGATCTG 372
 Db 358 ATGAGAAAGAGAGCTTGTCTAAAGATATCTCAAAAAGAGATGATGCAACCAAGATCTG 417
 Oy 417 ACTGAGGCACTAGAGTAGAGATTAAGATGCAATTTGAAAAATTGAGCAAGAGAGCTGTA 432
 Db 418 ACTGAGGCACTAGAGTAGAGATTAAGATGCAATTTGAAAAATTGAGCAAGAGAGCTGTA 477
 Oy 478 AAGTACAGAGCAACAGATTTGTAACCGGCTATTAAAGCTTAAAGCTTAAAGCTTAAAG 492
 Db 493 GTTGTAGAGCACTTCTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552
 Oy 552 GTTGTAGAGCACTTCTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
 Db 538 GTTGTAGAGCACTTCTCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597

Db	Accession	Version	Source	Organism	Reference	Authors	Title	Journal	Comment
Yy	553	GTGTCGCTCTGCTTCGACGAAGATATGAGACTCCCTTACTTTGGGCTCCA	603						
Db	598	GTGATGCTCTTCGATCAGACAGATATGAGACTCACTTACTTTGGAGCTCCA	648						
RESULT 10									
LOCUS	BE186786/c								
DEFINITION	946012C08.X1 946 - tassels primordium prepared by Schmidt Lab								
ACCESSION	BE186786								
VERSION	BE186786.1								
KEYWORDS	EST.								
SOURCE	Zea mays.								
ORGANISM	Zea mays.								
REFERENCE	Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC								
AUTHORS	Maize ESTs from various cDNA libraries sequenced at Stanford University								
TITLE	Unpublished (1999)								
JOURNAL	Contact: Walbot V								
COMMENT	Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 946012, row: C column: 08.								
FEATURES									
Source	Location/Qualifiers								
	1..467								
	/organism="Zea mays"								
	/cultivar="OH43"								
	/db_xref="taxon:4577"								
	/clone_lib="946 - tassels primordium prepared by Schmidt lab"								
	/tissue_type="tassels"								
	/dev_stage="just after the transition from vegetative to inflorescence development"								
	/lab_host="XLOLR"								
	/note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI; Site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Starfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."								
BASE COUNT	108 a 110 c 96 g 153 t								
ORIGIN									
Query Match	32.4%; Score 448; DB 9; Length 467;								
Best Local Similarity	98.7%; Pred. No. 3.8e-85;								
Matches	462; Conservative 0; Mismatches 5; Indels 1; Gaps 1;								
Yy	838	AGATATCAAAATTCCTGAGGACTGGCCTTACCAAGAAGCTCGACGCTTGTCAAGAGCCT	897						
Db	467	AGATATCAAAATTCCTGAGGACTGGCCTTACCAAGAAGCTCGACGCTTGTCAAGAGCCT	408						
Yy	898	AATGCACATTGGATATTCCTGAGCTAAATGAGATGACCTGATGAGAGGGTCTCTATA	957						
Db	407	AATGCACATTGGATATTCCTGAGCTAAATGAGATGACCTGATGAGAGGGTCTCTATA	348						
Yy	958	AGTTTCTGTGTAAGAATGTTTCAACAGAGATCGGGTGACAAAGCCATAGAGAG	1017						
Db	347	AGTTTCTGTGTAAGAATGTTTCAATGAGATCGGGTGACAAAGCCATAGAGAG	288						
Yy	1018	ATCAATCTGCCAGAAATTAATCGTCGGAAGAGAGACTCGAGTCTTTTCAAGCCAAT	1077						
Db	287	ATCAATCTGCCAGAAATTAATCGTCGGAAGAGAGACTCGAGTCTTTTCAAGCCAAT	228						
Yy	1078	GCCACACATCAGCACCCTTAAACGAAGAGACTTGGATTAACGAAGAGGACACT	1137						
Db	227	GCCACACATCAGCACCCTTAAACGAAGAGAGACTTGGATTAACGAAGAGGACACT	168						

QY	1138	GCACAAAGAAACAAAGCGCTGGTGGAAACAAACAAATTAATCTTGAGATCGTTGATGACAA	1197
Db	167	GCACAAAGAAACAAAGCGCTGGTGGAAACAAACAAATTAATCTTGAGATCGTTGATGACAA	108
QY	1198	CTACGACTACGAAGACAGCGCGTGGCGATCGACTTGCCTTAGATTATTAATCTCCCTGTT	1257
Db	107	CTACGACTACGAAGACAGCGCGTGGCGATCGACTTGCCTTAGATTATTAATCTCCCTGTT	49
QY	1258	TTTAACTCAGAGCTTTGGTAAAGTTGCGCCCATCTTCACGCTGGGGTA	1305
Db	48	TTTAACTCAGAGCTTTGGTAAAGTTGCGCCCATCTTCACGCTGGGGTA	1
RESULT 11			
LOCUS	AW774700	733 bp	mRNA
DEFINITION	EST333851 KV3 Medicago truncatula cDNA clone pKV3-23L21, mRNA		
ACCESSION	AW774700		
VERSION	AW774700.1	GI:7718617	
KEYWORDS	EST.		
SOURCE	barrel medic.		
ORGANISM	Medicago truncatula		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.		
REFERENCE	1 (bases 1 to 733)		
AUTHORS	Vandenbosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.		
TITLE	ESTs from roots of Medicago truncatula after Rhizobium inoculation		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Vandenbosch K Department of Biology Texas A&M University College Station, TX 77843-3258, USA Tel: 409 845 7707 Fax: 409 845 2891 Email: kate@mail.bio.tamu.edu Texas A&M EST name:T258119e TIGR sequence name:MTEBE71TK More information is available at: . http://chryslie.tamu.edu/medicago Seq primer: SKmod (CTA GAA CTA gta gta GAT CC).		
FEATURES			
source	location/qualifiers		
	1..733		
	/organism="Medicago truncatula"		
	/cultivar="genotype A17"		
	/db_xref="taxon:3880"		
	/clone="pKV3-23L21"		
	/clone_lib="KV3"		
	/tissue_type="Seedling roots"		
	/dev_stage="3 days post-inoculation with Sinorhizobium meliloti"		
	/lab_host="E. coli strain XLOLR"		
	/note="Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stragene and packaged using GigaPack ITI Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."		
BASE COUNT	222 a	125 c	183 g
ORIGIN	203 t		
	30.4%: Score 420.4; DB:9; Length 733;		
	Best Local Similarity 75.4%; Pred.No.2.8e-79;		
	Matches 523; Conservative 0; Mismatches 171; Indels 0; Gaps 0;		
36	GATGGGATCAAGGGTTTGACGAAACTGCTGGCGACAAATCGCCCAAGGCGATCAAGCA	95	

```

Db 40 GATGGGTATTAAGGCTTTAAGCAAGCTTTAGCTGATATATGCTCCCAATGATGAAGA 99
QY 96 GCAGAGTTTCAGAGACTTCTGGCCCAAAATGCGCTGCACCCAGCATGACATATA 155
Db 100 GAACAAATTCGAACTTCTTTGGGCTTAAGATGCTGTTGATGATGATGACATTTA 159
QY 156 CCAGTTCTGATTTAGTTGGAGGAGGATGGAAGCTTCACAAATGAAGCTGGTGA 215
Db 160 CAGTTTCTTATTTGTTGGGAGAGATGGAAGCTTGAAGTGTGACATGAACTGGTGA 219
QY 216 ACTCACTAGTCATTTGCAAGGAATGTTCAACCGCAATTAAGATTACTGGAAGCGGGAAT 275
Db 220 ACTACTGTCATTTGCAAGGAATGTTGGCGGACATACAGACTCTGAAGCCGGGAT 279
QY 276 CAAGCCAGTTTATGTTTGTGATGCGAACGCTTCGATATGAAGAAACAGAGCTTGCTAA 335
Db 280 GAAGCCAGATATGTTTGTGATGAGGAGCCACCGGAGATGAAGATCAAGAGCTGAAGAA 339
QY 336 AAGATATCAAAAAAGAGATGATGCAACCAAGATCTGACTGAGGAGTAGAGTAGAGA 395
Db 340 AGCTCTCTCAAGAGAGAGCTGAGGAGCTACCGCGGTTTGACAGAGAGCTTAGAGGCTGACAA 399
QY 396 TAAAGATGCGATTGAAAAATTTGAGCAGAGAGACTGTAAAGTTCACAGAGCAACACAGA 455
Db 400 TAAGGAAGATTTGAAAAATTCAGTAAACGACAGTAGAAGTGACAAAGCAATATATGA 459
QY 456 AGATGTTAAAGCGCTATTAGACTTATGAGGCTTCTGTTGATGAGGACCTTCTGAGAC 515
Db 460 CCAGTCGCAAAAACCTTTTGAACATGACATGAGGAGTGCCTGTTGAGGACACCTCGAAGAC 519
QY 516 AGAAGAGATGTCAGCCCTTGTGATTAACAGATAGAGTTCGCTGTGCTTGAAGA 575
Db 520 AGAGGCTGATGTCGTCGACCTTTCGAAGAGCTGGAAGAGTATGCTGAGGCTTGAAGA 579
QY 576 TATGAGCTCCCTTACTTTTGGGCTCCAGGCTTCTGCTCAATTAATGATCCAAAGTTC 635
Db 580 CATGATTCCTTAACCTTTGAGAGCTCTTAAGTTCTTCCCATTAATGATGATCCAAAGTTC 639
QY 636 CAGAAGATATCCGTCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 695
Db 640 AAGAGATATCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
QY 696 CATGACACGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
Db 700 CCTGACCAATTTATGACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733

RESULT 12
Bg263233 539 bp mRNA linear EST 16-FEB-2001
LOCUS Bg263233
DEFINITION WHE2339_F04.L072S Wheat pre-anthesis spike cDNA library Triticum
ACCESSION Bg263233
VERSION Bg263233.1 GI:12865185
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Triticum.
1 (bases 1 to 539)
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsiao,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
CONTACT: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 510595773
Fax: 510595818

```

```

FEATURES
source
1..539
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE239.F04.L07"
/clone_lib="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: lambda uni-ZAP XR, excised phagemid;
site.1: EcoRI; site.2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give pluscript
phagemids in the T7 Close lab (Choi, Close, Fenton) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT 163 a 108 c 141 g 126 t 1 others
ORIGIN
Query Match: 29.8%; Score 411.6; DB 10; Length 539;
Best Local Similarity 85.2%; Pred. No. 2,1e-77;
Matches 459; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
112 TACTTCGCGCCGCAAAATCGCCGTCGAGCCGACGATGATATACGATTCCTGATGTA 171
Db 1 TACTTCGCGCGCCGCGATCGCCGTCGAGCCGACGATGATATACGATTCCTGATGTA 60
QY 172 GTTGAAGGACGAGCATGCAATCTTCAAAATGAAGCTGTGAAGCACTATGCTATTG 231
Db 61 GTTGAAGGACGAGCATGCAATCTTCAAAATGAAGCTGTGAAGCACTATGCTATTG 120
QY 232 CAAGAAATGTTCAACCGGACATAGATTAAGTGAAGCGGAAATCAAGCCGTTTATGTT 291
Db 121 CAAGCAATGTTCAACCGGACATAGATTAAGTGAAGCGGAAATCAAGCCGTTTATGTT 180
QY 292 TTTGATGGCAAGCCTCTGATATGACAAACAGAGCTTCTTAAAGATCTCAAAAAG 351
Db 181 TTTGATGGCAAGCCTCTGATATGACAAACAGAGCTTCTTAAAGATCTCAAAAAG 240
QY 352 GATGATGCAACCAAGATGACAGCTGAGGAGAGTGAAGTGAAGATGATGGA 411
Db 241 AATGAGAGCAAGAGAGCTGTAAGGTCACAAAGGACACAGAGATGTAACGGCTA 300
QY 412 AATTTGAGCAAGAGAGCTGTAAGGTCACAAAGGACACAGAGATGTAACGGCTA 471
Db 301 AATTTGAGCAAGAGAGCTGTAAGGTCACAGAGGACACAGATGTAAGGCTGTA 360
QY 472 TTAAGACTTATGCGGCTTCTGTTGATGAGGACCTTCTGAGAGAGAGATGGA 531
Db 361 CTAAAGACTGAGGCTGTTCTGTTGATGAGGCTTCTGAGAGAGAGATGGA 420
QY 532 GCCCTTTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
Db 421 GCCCTTTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 592 TTTGGGCTCCACGCTTCTGTCATTTAATGATGATGATGATGATGATGATGATGATGATGAT 650
Db 481 TTTGGGCTCCACGCTTCTGTCATTTAATGATGATGATGATGATGATGATGATGATGATGAT 539

RESULT 13
AW562517/c 475 bp mRNA linear EST 10-MAR-2000
LOCUS AW562517
DEFINITION 66005H06.X2 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.

```

ACCESSION AM562517
VERSION AM562517.1 GI:7216395
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 475)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
CONTACT: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
FEATURES
source
1. 475
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 124 a 110 c 94 g 147 t
ORIGIN
Query Match 29.7%; Score 409.8; DB 9; Length 475;
Best Local Similarity 98.2%; Pred. No. 5.1e-77;
Matches 425; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 931 ACAGCAACCTGATGAGGAGGCTCATAGTTCCGTGTAAGAATATGTTTCACAGAA 990
DB 475 ACTGCACCTGATGAGGAGGCTCATAGTTCCGTGTAAGAATATGTTTCATGAA 416
QY 991 GATCGGGTGAACAAGGCCATAGAGAAGATCAATCTGCCAAGATTAATCTGCGCAAGA 1050
DB 415 GATCGGGTGAACAAGGCCATAGAGAAGATCAATCTGCCAAGATTAATCTGCGCAAGA 356
QY 1051 AGACTGAGTCCCTTTTCAAGCAACTGCCACACATCGCACCGCTTAACGCGAAGAG 1110
DB 355 AGACTGAGTCCCTTTTCAAGCAACTGCCACACATCGCACCGCTTAACGCGAAGAG 296
QY 1111 ACTTCGATTAACAACAAGCAGCAGCTGCGAACAAGAAACAAGAGCTGTGGAAGAAG 1170
DB 295 ACTTCGATTAACAACAAGCAGCAGCTGCGAACAAGAAACAAGAGCTGTGGAAGAAG 236
QY 1171 AATAATCTTGATGCTGTGATGACACTACAGACTACGAACACGCGTGCCTATAC 1230
DB 235 AATAATCTTGATGCTGTGATGACACTACAGACTACGAACACGCGTGCCTATAC 176
QY 1231 TTGCGTAGATTATTAACCTCCCTGTTTAACTCAGAGCTTGTGAAGTTCGCCCATG 1290
DB 175 TTGCGTAGATTATTAACCTCCCTGTTTAACTCAGAGCTTGTGAAGTTCGCCCATG 116
QY 1291 TTTCAGCTGGGGTAAGTACTGTGTGGAAGACTTGGTACCAGTAACAACCTT 1330
DB 115 TTTCAGCTGGGGTAAGTACTGTGTGGAAGACTTGGTACCAGT-ACAGAACTT 57
QY 1351 ATGCGTGTGTTT 1363
DB 56 ATGCGTGTGTTT 44

RESULT 14
AM288784 414 bp mRNA linear EST 16-JAN-2000
LOCUS AM288784
DEFINITION 707010C02.x5 707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA; mRNA sequence.
ACCESSION AM288784
VERSION AM288784.1 GI:6695706
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 414)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
CONTACT: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: C column: 02.
FEATURES
source
1. 414
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/21/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT 120 a 82 c 101 g 109 t 2 others
ORIGIN
Query Match 29.1%; Score 402.2; DB 9; Length 414;
Best Local Similarity 98.8%; Pred. No. 2.1e-75;
Matches 404; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 601 CCACGGTTCCTTCGTCATTTATGATCCAAAGTTCCAGAAATACCTGTGATGATTTG 660
DB 6 CCACGGTTCCTTCGTCATTTATGATCCAAAGTTCCAGAAATACCTGTGATGATTTG 65
QY 661 GATGTGCCCAAGCTTTTGGAGAGAGCTTGAACATCACCATGACACGATTCATGATTTGTC 720
DB 66 GATGTGCCCAAGCTTTTGGAGAGAGCTTGAACATCACCATGACACGATTCATGATTTGTC 125
QY 721 ATCTGTGTGATGTGACTATTTGTATAGCATCAAGATATCGGGGGCAACAGCTCTG 780
DB 126 ATCTGTGTGATGTGACTATTTGTATAGCATCAAGATATCGGGGGCAACAGCTCTG 185
QY 781 AACTATTTCGTCACATGGTGTCCATAGAAAGCATCTGGAGAAATCTTAATAAAGACAGA 840
DB 186 AACTATTTCGTCACATGGTGTCCATAGAAAGCATCTGGAGAAATCTTAATAAAGACAGA 245
QY 841 TATCAATTCCTGAGAGCTGCGCTTACACAGAGCTGCAGCGCTTGTTCAGGACCTAAT 900
DB 246 TATCAATTCCTGAGAGCTGCGCTTACACAGAGCTGCAGCGCTTGTTCAGGACCTAAT 305
QY 901 GTCACATTTGATATTCCTAGCTAAATATGAGTACCTGATGAGAGAGGCTCATAGT 960
DB 306 GTCACATTTGATATTCCTAGCTAAATATGAGTACCTGATGAGAGAGGCTCATAGT 365

```

QY 961 TTCTCGTAAAGATATGTTTCACGACATGCGGTGACAAAGGCCA 1009
    |||||||
Db 366 TTCCGCGTAAAGATATGTTTCACGACATGCGGTGACAAAGGCCA 414

RESULT 15
LOCUS A1065689/c
DEFINITION ag91f12.x1 maize inflorescence immature ear library Zea mays CDNA
ACCESSION A1065689
VERSION A1065689
KEYWORDS A1065689.1 GI:3341096
SOURCE EST.
ORGANISM Zea mays.
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 456)
            Schutz,K., de la Bastide,M., Gnoj,L., Habermann,K., Huang,E.N.,
            Parnell,L.D., Dedhia,N., Martensen,R. and McCombie,W.R.
            Expressed sequence tags from Z. mays
            Unpublished (1998)
            Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mcombie@cshl.org
            Plate: ag91 row: f column: 12
            Seq primer: M13 forward universal -21
            High quality sequence stop: 456.
            Location/Qualifiers
                1..456
                /organism="Zea mays"
                /cultivar="B73"
                /db_xref="taxon:457"
                /clone="ag91f12"
                /clone_lib="maize inflorescence immature ear library"
                /sex="female"
                /tissue_type="immature ear"
                /note="Vector: pBLUESCRIPT SK+ (X52325); Site.1: XhoI;
                Site.2: EcoRI; This library is described in Schmidt, Hake,
                et al., (1993) Plant Cell 5:729-737. cDNAs are
                directionally cloned into the XhoI and EcoRI sites. XhoI
                is near the polyA tail. Most reads from this library are
                3' in direction. Additional information on this library as
                well as ftp access to all sequences can be found at
                http://www.cshl.org/maizegenome"

BASE COUNT 112 a 104 c 86 g 152 t 2 others
ORIGIN

Query Match 28.7%; Score 397; DB 9; Length 456;
Best Local Similarity 91.9%; Pred. No. 2.7e-74;
Matches 418; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 909 GGATATTCCTGAGCTAAATGCGACTGCACCTGATGAGAGGAGTTCATAGTTTCTGCT 968
    |||||||
Db 456 GGATATTCCTGAGCTAAATGCGACTGCACCTGATGAGAGGAGTTCATAGTTTCTGCT 397

QY 969 AAAAGATATGTTTCACGACATGCGGTGACAAAGGCCATAGAGAGATCAATCTGC 1028
    |||||||
Db 396 AAAAGATATGTTTCACGACATGCGGTGACAAAGGCCATAGAGAGATCAATCTGC 337

QY 1029 CAAGATATATGCTGCGAAGAGACTGAGTCTTTTTCAGCCAACTGCCACCATC 1088
    |||||||
Db 336 CAAGATATATGCTGCGAAGAGACTGAGTCTTTTTCAGCCAACTGCCACCATC 277

QY 1089 AGCACCGCTAAAGCGAAGGAGACTTGGATTAACCAAGCAGCAGCTCGAACAAGAA 1148
    |||||||
Db 276 AGCACCGCTAAAGCGAAGGAGACTTGGATTAACCAAGCAGCAGCTCGAACAAGAA 217

```

```

QY 1149 AACAAAGCTGTGGAAAGAGAAATATCTTGATGCTTGATGACACTACGACTACG 1208
    |||||||
Db 216 AACAAAGCTGTGGAAAGAGAAATATCTTGATGCTTGATGACACTACGACTACG 157

QY 1209 AAAGCAGCGGTGCGGTGATCAGTTCCTGTTAGATTAATTAACCTCGTTTAACTCAAG 1268
    |||||||
Db 156 AAAGCAGCGGTGCGGTGATCAGTTCCTGTTAGATTAATTAACCTCGTTTAACTCAAG 97

QY 1269 CTTTGTAAAGTTCGCCCATGTTTCAAGCTGCGGTAAGTTAGTTGTGTTGAAGAGATT 1328
    |||||||
Db 96 TGAAGAGATNGGTGTACCAAGTAACAAACTTATGCGTGTGTTTATGTTGAAGAGATT 37

QY 1329 GGTGTACCAAGTATACAAACTTATGCGTGTGTTT 1363
    |||||||
Db 36 GGTGTACCAAGTATACAAACTTATGCGTGTGTTT 2

Search completed: November 5, 2002, 13:36:58
Job time : 1195.61 secs

```


QY	319	TTACGTGGAAAGCGGGAATCAACCCAGTTATGTTTTGATGGCAAGCCTCTGATATGAAG	378
QY	307	TTACTGGAAAGGGGGAATCAACCCAGTTATGTTTTGATGGCAAGCCTCTGATATGAAG	366
QY	379	AAACAAAGAAGCTTGGCTAAAAAGATTACTCAAAAAAGATGATCAACCCAAAGATCGATCGAG	438
Db	367	AAACAAAGAAGCTTGGCTAAAAAGATTACTCAAAAAAGATGATCAACCCAAAGATCTGACTGAG	426
QY	439	GCAGTAGAGGTGAGGAGATAAAGATGCGATTTGAAAAAATTGAGCAAGAGCACTGTAAAGTTC	498
Db	427	GCAGTAGAGGTGAGGAGATAAAGATGCGATTTGAAAAAATTGAGCAAGAGCACTGTAAAGTTC	486
QY	499	ACAAGGCAACCAACGAAAGATTTGATAACGACTATTAAGACTTTGCGGGTCTCTGTTGTA	558
Db	487	ACAAGGCAACCAACGAAAGATTTGATAACGACTATTTAAGACTTTGCGGGTCTCTGTTGTA	546
QY	559	GAGGCAACCTTCTGAAAGCAAGAGCAAGATGTGCAGCCCTTGCATTAAGCATTAAGGTTTC	618
Db	547	GAGGCAACCTTCTGAAAGCAAGAGCAAGATGTGCAGCCCTTGCATTAAGCATTAAGGTTTC	606
QY	619	GCTGTTGCTTCAGAGAATAAAGACTCCCTTACTTTTGGGGCTCCACGTTCCCTGCTCAT	678
Db	607	GCTGTTGCTTCAGAGAATAAAGACTCCCTTACTTTTGGGGCTCCACGTTCCCTGCTCAT	666
QY	679	TTAATGATTCGAAGTTCCTCAAGAAATACCTGTATGCAATTTGATGTTGCCAAGGTTTG	738
Db	667	TTAATGATTCGAAGTTCCTCAAGAAATACCTGTATGGAATTTGATGTTGCCAAGGTTTGG	726
QY	739	GAGGAGCTTGAAGTCAACCATGAGACAGTTCATTTGATTTGGCATCCCTGTGTGATGTAC	798
Db	727	GAGGAGCTTGAAGTCAACCATGAGACAGTTCATTTGATTTGGCATCCCTGTGTGATGTAC	786
QY	799	TATTGTGATACCATCAAAAGGTATTCGGGGGCAACAGCTCTGAAACTTATTTGTCACAT	858
Db	787	TATTGTGATACCATCAAAAGGTATTCGGGGGCAACAGCTCTGAAACTTATTTGTCACAT	846
QY	859	GGGTCCATAGGAAGCAATCTTGGAGAACTTAAATAAGACAGATATCAAAATTTCTCAGAC	918
Db	847	GGGTCCATAGGAAGCAATCTTGGAGAACTTAAATAAGACAGATATCAAAATTTCTCAGAC	906
QY	919	TGGCCTTACCAAGAGGCTCGACGCTTGTTCAAAGACCTTAAGTCACTTGGATATTTCT	978
Db	907	TGGCCTTACCAAGAGGCTCGACGCTTGTTCAAAGACCTTAATGTCACTTGGATATTTCT	966
QY	979	GAGCTAAATGAGCTGCACCTGATGAGGAGGTTCTCATAGTTTCCGTGTAAAGATAT	1038
Db	967	GAGCTAAATGAGCTGCACCTGATGAGGAGGTTCTCATAGTTTCCGTGTAAAGATAT	1026
QY	1039	GGTTTCATGAAGTTCGGGGTGACAAAGGCCATTAAGAAATCAAAATCTGCCAAGATATA	1098
Db	1027	GGTTTCATGAAGTTCGGGGTGACAAAGGCCATTAAGAAATCAAAATCTGCCAAGATATA	1086
QY	1099	TGCTGCGAAGGAAACATCGAGTCCCTTTTTCACCAAGTCCACCAACATCACACCGGTA	1158
Db	1087	TGCTGCGAAGGAAACATCGAGTCCCTTTTTCACCAAGTCCACCAACATCACACCGGTA	1146
QY	1159	AAACGGAAGAGACTTCGGATTAACCAAGCAGCAGCTGCGACACAGAAACAAAGAGCT	1218
Db	1147	AAACGGAAGAGAGACTTCGGATTAACCAAGCAGCAGCTGCGACACAGAAACAAAGAGCT	1206
QY	1219	GGTGGAAAGAAATATATCTTGGATGTGTTGATGTCAACTACGACTACGAAGCAGCGG	1278
Db	1207	GGTGGAAAGAAATATATCTTGGATGTGTTGATGTCAACTACGACTACGAAGCAGCGG	1266
QY	1279	TGGGATGATCACTTCGCGCTAGATTATTTTAACGCCCTGTTTAAGTCAGACGTTGGGTAA	1338
Db	1267	TGGGATGATCACTTCGCGCTAGATTATTTTAACGCCCTGTTTAAGTCAGAGCTTGGGTAA	1326
QY	1339	AGTTTGCATGTTTCAAGCTGGGGGTAGTTAGTTGTGTTGAAGAGATTTGGTGTACCAA	1398
Db	1327	AGTTTGCATGTTTCAAGCTGGGGGTAGTTAGTTGTGTTGAAGAGATTTGGTGTACCAA	1386
QY	1399	GTAACAACACTATCGCTGTTTTTACTCTCTTCTTGAAGATGTATGCGCAAGTATAA	1458

	D	b	1387	GTAACAAACTATCGCTTTTTCATTGCTCTTGAGGTAAAAA	A	AAAAAAAAAAAAAAAAAAA	1465	1446
OY	1459	AAAAAAAAAAAAAAAAAAA	1475					
D	1447	AAAAAAAAAAAAAAAAAAA	1463					
R	E	S	I	L	T			
RESULT 3								
LOCUS	ARI52404		1541 bp	DNA		linear	PAT 08-AUG-2001	
DEFINITION	Sequence 3 from patent US 6232527.							
ACCESSION	ARIS2404							
KEYWORDS	ARIS2404.1 GI:15118454							
SOURCE	.							
ORGANISM	Unknown..							
REFERENCE	Unclassified.							
AUTHORS	1 (bases 1 to 1541)							
TITLE	Mahajan,P.B.							
JOURNAL	Maize Rad2/FEN-1 orthologues and uses thereof							
FEATURES	Patent: US 6232527-A 3 15-MAY-2001;							
	Location/Qualifiers							
	source							
BASE COUNT	473 a	308 c	377 g	383 t				
ORIGIN	/organism="unknown"							
Query Match		94.8%;	Score 1401.2;	DB 6; Length 1541;				
Best Local Similarity		99.1%; Pred. No. 0;						
Matches 1409; Conservative		0; Mismatches 13; Indels 0; Gaps 0;						
OY	33	GGGGTTTTCTGGGCCACCCTCCGCAGC	C	CGCCGCCCCGCAACCCGACAAGCCC	G	CAGA	92	
D	15	GGCGTTTTCTTGCGCACCTCCG	C	CTAGCCGCCGCCACCCGCAAGCCG	C	GAGA	74	
OY	93	CGAGATGGGCATCAAAGG	G	TTTGACGAACATGCTG	SGCGACAATGCGCCAAGCGCATGAA	152		
D	75	CGAGATGGGCATCAACAGG	G	TTTGACGAACATGCTG	SGCGACAATGCGCCAAGCGCATGAA	134		
OY	153	GGAGCAGAAGTTCGAGAG	C	TACTTGGCCGCAAAATCGCCGTCGACGC	CAGCATBAGCAT	212		
D	135	GGAGCAGAAGTTCGAGAG	C	TACTTGGCCGCAAAATCGCCGTCGACGC	CAGCATBAGCAT	194		
OY	213	CTAACCACTCTCGATAGTAG	TGTAAGTGAAGGCAAGGCATGGAAACTC	TCACAAATAAGTGAAGCTGG	272			
D	195	ATACCACTTCCATGATTG	TGTAAGTGAAGGCAAGGCATGGAAACTC	TCACAAATAAGTGAAGCTGG	254			
OY	273	TGAAGTCACTACTCATTT	TGCAAGAAATGTTCAACCGGCAATTAAGATPAC	TGSAAGCGGG	332			
D	255	TGAAGTCACTACTCATTT	TGCAAGAAATGTTCAACCGGCAATTAAGATPAC	TGSAAGCGGG	314			
OY	333	AATCAAGCCAGTTAT	TGTTTTGATGGCAAGCTCTCGATATGAAGAAACAAGACTTGC	392				
D	315	AATCAAGCCAGTTAT	TGTTTTGATGGCAAGCTCTCGATATGAAGAAACAAGACTTGC	374				
OY	393	TAAAGATAC	TCAAAAAGAGATGATGCACAACCAAGATGTGCTGAGCGAGTAGAGTAGG	452				
D	375	TAAAGATAC	TCAAAAAGAGATGATGCACAACCAAGATGTGCTGAGCGAGTAGAGTAGG	434				
OY	453	AGATAAAGATGGGAT	TGAAAAATTGAGCAAGAGCACTTAAGGTCAACAAAGCAACAA	512				
D	435	AGATAAAGATGGGAT	TGAAAAATTGAGCAAGAGCACTTAAGGTCAACAAAGCAACAA	494				
OY	513	CGAAGATTGTAACGAC	TATTBAAGCTTATGGGGGTTCTGTGTGAGAGCACTTCTGA	572				
D	495	CGAAGATTGTAACGAC	TATTBAAGCTTATGGGGGTTCTGTGTGAGAGCACTTCTGA	554				
OY	573	AGCACAACGAGANTG	AGCCCTTTGGATPAAGATPAAGSTGTCCGTGTGCTTACA	632				
D	555	AGCACAACGAGANTG	AGCCCTTTGGATPAAGATPAAGSTGTCCGTGTGCTTACA	614				
OY	633	AGATAAAGACTCC	TTACTTTTGGGGCGCACGGTCTCTTGATTTAAATGATTCGAAG	692				

BASE	COUNT	441 a	269 c	346 g	325 t	ORIGIN
Query Match	90.3%:	Score 1334.2:	DB 6:	Length 1381:		
Best Local Similarity	98.7%:	Pred. No. 0:				
Matches 1345:	Conservative	0:	Mismatches 18:	Indels 0:	Gaps	
OY	61	CGCGGCCCCGACCCCGCCACAGCCGCCGCGAGACGAGATGGGCATCMAGGGTTTGACGAA	120			
Db	1	CGACCCACAGGGTCCGGCCACAGCCGCCGCGAGAGATGGGCATCAAGGGTTTGACGAA	60			
OY	121	CTGCTGGCGACAAATGCGCCCAAGCGGATGAAGAGCAGAAATTTGAAAGACTACTTGGC	180			
Db	61	CTGCTGGCGGACAAATGCGCCCAAGGGGATGAAGAGCAGAAATTTGAAAGACTACTTGGC	120			
OY	181	CGCAAAATGCGCCGCGACGCGCACATGAGCATCTACCATTTCTGATAGTAGTTGAAG	240			
Db	121	CGCAAAATGCGCCGCGCACGCGCACATGAGCATCTACCATTTCTGATAGTAGTTGAAG	180			
OY	241	ACAGGCATGGAACCTCTCACAAATGAAGCTGGTGAAGTCACTAGTCATTTGCCAAGAA	300			
Db	181	ACAGGCATGGAACCTCTCACAAATGAAGCTGGTGAAGTCACTAGTCATTTGCCAAGAA	240			
OY	301	TTTCAACCGGACAAATACATTAATCTGGAAGGGGAATCAACCCGTTATGTTTGAATGC	360			
Db	241	TTTCAACCGGACAAATACATTAATCTGGAAGGGGAATCAACCCGTTATGTTTGAATGC	300			
OY	361	AAGCCTCCTGATATGAAGAAACAAGAACTTGTCAAAAGTACTCAAAAAGAGATGATCA	420			
Db	301	AAGCCTCCTGATATGAAGAAACAAGAACTTGTCAAAAGTACTCAAAAAGAGATGATCA	360			
OY	421	ACCAAAAGTCTGACTGAGCGAGTAGAGTAGAGGATTAAGATGGCATTTGAAAATTTGAGC	480			
Db	361	ACCAAAAGTCTGACTGAGCGAGTAGAGTAGAGGATTAAGATGGCATTTGAAAATTTGAGC	420			
OY	481	AAGAGACTGTAAAGGTGACACAAAGCAACAAACGAAGATTTGAAACGACTATTAAGATT	540			
Db	421	AAGAGACTGTAAAGGTGACACAAAGCAACAAACGAAGATTTGAAACGACTATTAAGATT	480			
OY	541	ATGGGGGTTCTGTTGTAGAGGCACTTCTGAAGCAGAAAGCAATGTGCAGCCCTTTGC	600			
Db	481	ATGGGGGTTCTGTTGTAGAGGCACTTCTGAAGCAGAAAGCAATGTGCAGCCCTTTGC	540			
OY	601	ATAAACGATTAAGGTTGCGGTGCTTGAAGATTAAGAGACGCCCTTACTTTGGGCT	660			
Db	541	ATAAACGATTAAGGTTGCGGTGCTTGAAGATTAAGAGATGAGACTCCCTTACTTTGGGCT	600			
OY	661	CCACGGTCTCTCTCATTTTAATGATCCAAAGTTCCAGAATAATACCTGTGATGAATTT	720			
Db	601	CCACGGTCTCTCTCATTTTAATGATCCAAAGTTCCAGAATAATACCTGTGATGAATTT	660			
OY	721	GATGTTGCCAAGTTTGGGAGAGCTTGAACCTACCATGGACAGTCAATGATTTGGC	780			
Db	661	GATGTTGCCAAGTTTGGGAGAGCTTGAACCTACCATGGACAGTCAATGATTTGGC	720			
OY	781	ATCCGTGTGATGATGACTATTGTGATAGCATCAAGGATGTGGGGGGCAACACCTGTG	840			
Db	721	ATCCGTGTGATGATGACTATTGTGATAGCATCAAGGATGTGGGGGGCAACACCTGTG	780			
OY	841	AAACTTAATTCGTCAACATGsbhTCANAGAAGCATTTGGAGAAATCTTAATAAGACAGA	900			
Db	781	AAACTTAATTCGTCAACATGsbhTCATGGAAGCATTTGGAGAAATCTTAATAAGACAGA	840			
OY	901	TATCAAAATTCCTGAGGACTGGCCTTACCAAGAGCTGAGGCTGTGTCAGAGGCTAT	960			
Db	841	TATCAAAATTCCTGAGGACTGGCCTTACCAAGAGCTGAGGCTGTGTCAGAGGCTAT	900			
OY	961	GTCACATTTGATATTCCTGAGCTAAATAGACACACCTGATAGAGAGGCTCTTAAGT	1020			
Db	901	GTCACATTTGATATTCCTGAGCTAAATAGACACACCTGATAGAGAGGCTCTTAAGT	960			
OY	1021	TTTCCTGTAAGATTAATGTTTTCATGAAGATCGGGTGCAAAAGCCATAGAGAAATC	1080			

Db 961 TTCTCGTGAAGATAGGTTTCAACGAGATCGGGTGACAAAGGCCATAGAGAATC 1020
Oy 1081 AATCTGCCAAGATTAATCTGCGACAGAGACTCGAGTCTTTTTCAGGCACTGCC 1140
Db 1021 AATCTGCCAAGATTAATCTGCGACAGAGACTCGAGTCTTTTTCAGGCACTGCC 1080
Oy 1141 ACCACATCGACACCGCTTAACGAGAGACTTGGATTAACAGCAAGGCACTGCC 1200
Db 1081 ACCACATCGACACCGCTTAACGAGAGACTTGGATTAACAGCAAGGCACTGCC 1140
Oy 1201 AACCAAAACAAAGGCTGGTGAAGAGAAATATCTTGATCGTGGATACACCTA 1260
Db 1141 AACCAAAACAAAGGCTGGTGAAGAGAAATATCTTGATCGTGGATACACCTA 1200
Oy 1261 CGACTACGAAAGCAGGCTGGCATGATCCTCGCCTAGATTATTAACCTCCCTGTTT 1320
Db 1201 CGACTACGAAAGCAGGCTGGCATGATCCTCGCCTAGATTATTAACCTCCCTGTTT 1260
Oy 1321 ACTCAGACTTGGTGAAGGTTGGCCATGTTCAAGCTGGGCTAGTGTGCTTG 1380
Db 1261 ACTCAGACTTGGTGAAGGTTGGCCATGTTCAAGCTGGGCTAGTGTGCTTG 1320
Oy 1381 AAGAGATTGGTGTACCACTACAAACTTATCGCTGTTT 1423
Db 1321 AAGAGATTGGTGTACCACTACAAACTTATCGCTGTTT 1363

RESULT 5
AB021666 1354 bp mRNA linear PLN 25-MAR-2000
LOCUS AB021666
DEFINITION Oryza sativa OsFEN-1 mRNA for FEN-1, complete cds.
ACCESSION AB021666
VERSION AB021666.2 GI:4587224
KEYWORDS FEN-1; endonuclease.
SOURCE Oryza sativa cDNA to mRNA.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1 (sites)
AUTHORS Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and
Sakaguchi,K.
TITLE Plant homologue of flap endonuclease-1: molecular cloning,
characterization, and evidence of expression in meristematic
tissues
JOURNAL Plant Mol. Biol. 42 (3), 415-427 (2000)
MEDLINE 20256470
AUTHORS 2 (bases 1 to 1354)
TITLE Kimura,S., Hashimoto,J. and Sakaguchi,K.
JOURNAL Direct Submission
SUBMITTED (21-DEC-1998) Seisuke Kimura, Science University of
Tokyo, Dept. of Applied Biological Science; 2641 Yamazaki, Noda,
Chiba 278-8510, Japan (E-mail:j5498703@nodai.ac.jp,
Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)
COMMENT On Apr 17, 1999 this sequence version replaced gi:4062866.
FEATURES
source
1..1354
/organism="Oryza sativa"
/db_xref="taxon:4530"
67..1209
/gene="OsFEN-1"
67..1209
/gene="OsFEN-1"
/note="OsFEN-1"
/note="endonuclease"
/codon_start=1
/product="FEN-1"
/protein_id="BAA36171.1"
/db_xref="gi:4587225"
/translation="MGIRGTRKILADNAPKAKMEKPESEGRRIAYDASMSYIFLI
VYGRGEMETLNEAGEVSHDGMENRIRILEAKIKYVYDGPPLKQELAKRY
SKREDATKELTVAVEGKDAIEFESKRTVTKQHNSEKRLTLKGVPAVEAPCEA
EAECALCINDVAVASEDMDSIFGAPRLRHLMDPSKRTIPVMEVAVKLEELE
LIMDOFIDLCITISGCDYSDISIKIGIGATALKLIRHGSIESILENTINRDRQIPEDWP

YOEARLEKPNVTLIDIPELKMNAPDEEGLVEFLVKNENFNDRVTKALEIKFAKNK
SSQRIESEFFKPVVSTSVPLKRDITSEKPTKAVANKRTKGAGKKK"
polyA_site 1354
/note="17 A nucleotides"
BASE COUNT 415 a 286 c 331 g 322 t
ORIGIN

Query Match 62.4%; Score 922.6; DB 8; Length 1354;
Best Local Similarity 86.1%; Pred. No. 1.7e-218;
Matches 1034; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

Oy 67 CCGCCACCCGCCACAGCCGCCGACAGACAGATGGGCATCAAGGTTTGAGCAACTCTG 126
Db 37 CCGCCGCTCGCGCATCCCTCGAGGCGAGATGGGCAATCAAGGTTTGAGCAACTCTG 96
Oy 127 CGGCAATATGCGCCCAAGAGCGATGAAGAGAGCAAGTTTCAGAGCTACTTCGCGCCAAA 186
Db 97 CGGCAATATGCGCCCAAGAGCGATGAAGAGAGCAAGTTTCAGAGCTACTTCGCGCGC 156
Oy 187 ATCGCCCTCGAGCGCCAGCATGATGATCACTTCTGATAGTATGATGAGACAGCGC 246
Db 157 ATCGCCCTCGAGCGCCAGCATGATGATCACTTCTGATAGTATGATGAGACAGCGC 216
Oy 247 ATGCAACTCTCACAAATAGAGCTGTGAGTCACTAGTCAATTTGCAAGAAATTTCAAC 306
Db 217 ATGCAAACTCTCACAAAGAGAGCGCGGTAACTACACATCTTTGCAAGATTTCAAC 276
Oy 307 CGGCAATATGATTAATCTGGAAGCGGGAATCAAGCCAGTTTATGTTGATGCAAGCCT 366
Db 277 CGGCAATATGATTAATCTGGAAGCGGGAATCAAGCCAGTTTATGTTGATGCAAGCCT 336
Oy 367 CCTGATTAAGAAACAAAGAACTGCTTAAGATTAATCAAGAGATGATGCAAGCAAA 426
Db 337 CCTGATTAAGAAACAAAGAACTGCTTAAGATTAATCAAGAGATGATGCAAGCAAA 396
Oy 427 GATCTGACTGAGGCGATGAGAGGTAAAGATGAGATGAGATGAGATGAGATGAG 486
Db 397 GATCTGACTGAGGCGATGAGAGGTAAAGATGAGATGAGATGAGATGAGATGAG 456
Oy 487 ACTGTAAGGTCACAGGCGACACACAGAGATGATGATGATGATGATGATGATGATG 546
Db 457 ACTGTAAGGTCACAGGCGACACACAGAGATGATGATGATGATGATGATGATGATG 516
Oy 547 GTTCTGTTGTAAGAGCGCTTCTGAAGCAGAGAGATGAGAGAGAGAGAGAGAGAG 606
Db 517 GTTCTGTTGTAAGAGCGCTTCTGAAGCAGAGAGATGAGAGAGAGAGAGAGAGAG 576
Oy 607 GATTAAGGTTCGCTGCTGCTTTCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 666
Db 577 GATTAAGGTTCGCTGCTGCTTTCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 636
Oy 667 TTCTCTGCTCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 726
Db 637 TTCTCTGCTCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 696
Oy 727 GCCAAGGTTTGGAGAGAGCTTGAACATGACAGCAGTTCATGATTTGATGATCCTG 786
Db 697 GCCAAGGTTTGGAGAGAGCTTGAACATGACAGCAGTTCATGATTTGATGATCCTG 756
Oy 787 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
Db 757 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 816
Oy 847 ATTCGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
Db 817 ATTCGTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 876
Oy 907 ATTCGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
Db 877 ATTCGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 936
Oy 967 TTGATATTCCTGAGCTAAATGAGCTGACCTGATGAGAGAGGCTTCATAGTTTCTCG 1026
Db 936 TTGATATTCCTGAGCTAAATGAGCTGACCTGATGAGAGAGGCTTCATAGTTTCTCG 1026

Db 1010 GTGAGCTAGGCTGATGTAAGAGGAGACTGGTGGCCCTTCATGTGTGGGAGAGAAGCAGTTTCAG 1069
Qy 1047 TGAATGAGGGGTGACAAAGGCCATAGAGAAGATCAAAATGTGCCAAGATATATCGTGC 1106
Db 1070 TGAATGAGGGGTGACAAAGGCCATAGAGAAGATCAAAATGTGCCAAGAGCCGTCAGAGCAGCA 1129
Qy 1107 AGGAAGACTCGAGTCTTTTTCAGAACCACTGCCACCA 1144
Db 1130 GGGCCGCTGATGATCTTTTTCAGAACGATGATGATCA 1167

RESULT 7

XU064563

LOCUS XU064563 1461 bp mRNA linear VRT 26-JUL-1999
DEFINITION Xenopus laevis 5' nuclease xFEN1A mRNA, complete cds.
ACCESSION U64563
VERSION U64563.1 GI:1490869

KEYWORDS

SOURCE African clawed frog.
ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;

Xenopodinae; Xenopus.
1 (bases 1 to 1461)
Bibikova, M., Wu, B., Chi, E., Kim, K.H., Trautman, J.K. and Carroll, D.

Characterization of FEN-1 from Xenopus laevis. cDNA cloning and
role in DNA metabolism

J. Biol. Chem. 273 (51), 34222-34229 (1998)
MEDLINE 99069415
PUBMED 9852084

2 (bases 1 to 1461)
Bibikova, M., Chi, E., Wu, B., Kim, K.-H. and Carroll, D.

Direct Submission
Submitted (19-JUL-1996) Biochemistry, U. of Utah Med. Sch., 50 N.

Journal Medical Dr., Salt Lake City, UT 84132, USA
LOCATION/Qualifiers

1. 1461
/organism="Xenopus laevis"
/db_xref="taxon:8355"

111. 1259
/function="Implicated in processing Okazaki fragments
during DNA replication and in base excision repair;
specifically removes single-stranded 5' sequence"

/note="Similar to FEN-1, Dnase1, WFL, RAD27, RTH1, and
rad2"

/codon_start=1
/product="5' nuclease xFEN1A"

/protein_id="AAB06176.1"

/db_xref="GI:1490870"

/translation="MGTHGLAKLADYAPAIKEHDIKSYFGRKRVAVDSMCYOFLL
AVRODNTLONEBEGTTHSLMGEFTIRNVEHGRKRVYVDFGKPKOMKSGELAKSE
RRABEKLEAEAGAEVENIEKFTKRLVTVQNHNECKLLTGLPYVADPCAE
ATCAALYAGKRVAAATEMDALFTGPRVLRHTLTAESAKRLIOGFHNRYVDIGI
THEQFVDCILGSDYCEETIRIGIPKRAIDLIROHRTIDIIINIDLKRYPYENMLH
KEAKHLEFEVVDITTELMKIGPDEGGLVAVAGKGFSEPRIRNGKAKLAKNROG
STGRLDDEFVTVSVSTFKREAESGSKAKKAKGTGPAKFKRKGK"

BASE COUNT 454 a 293 c 365 g 349 t

ORIGIN

Query Match 21.0%; Score 310.2; DB 5; Length 1461;
Best Local Similarity 55.8%; Pred. No. 3.1e-66;

Matches 635; Conservative 0; Mismatches 498; Indels 6; Gaps 2;

Qy 97 ATGGGATCAAGGTTTGCAAGAACTGCGGGGAGCAATGGGCCCAAGGGGATGAGAG 156
Db 111 ATGGGATTCACGGTTTGGCCAACTTATCGCAGATGTGGCACTGCGAGCTATCAAGAG 170

Qy 157 CAGAACTTCGAGACTACTTCGGCGCAAAATCGCCGTGAGCGCAGCATGAGCATCTAC 216
Db 171 CATGATATATAAAGTACTTGTGTAAGAGGCGGCGTGCATGCTCATGATTTAT 230

Qy 217 CAGTTCCTATAGTGTGAGAGCAGCAGCATGCAAACTCCACAAATGAGCTGGTGAA 276
Db 111 CAGTTCCTATAGTGTGAGAGCAGCAGCATGCAAACTCCACAAATGAGCTGGTGAA 276

Qy 217 CAGTTCCTATAGTGTGAGAGCAGCAGCATGCAAACTCCACAAATGAGCTGGTGAA 276
Db 111 CAGTTCCTATAGTGTGAGAGCAGCAGCATGCAAACTCCACAAATGAGCTGGTGAA 276

Db 231 CAGTTTCTTATGAGTGAAGACAGATGGCA--ACAGCTGCAAAAAGAGAGGTGAA 287
Qy 277 GTCACTAGTCAATTTGCAAGCAATGTTCAACCGGACATTAAGATTACGNAAGGGGAAATC 336
Db 288 ACCCAACCCATCTTATGAGGATGTTTACCGTACATTTGGCATGTGAGAGCATGGGCATC 347
Qy 337 AAGCCAGTTTATGTTTGGATGSCAAGCCTCTGTGATATGAGAAACAGACATCTGTA 396
Db 348 AAGCCAGTTTATGTTTGGATGSCAAGCCTCTGTGATATGAGAAACAGACATCTGTA 407
Qy 397 AGATACCTCAAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456
Db 408 CCAATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 467
Qy 457 AAGATGCGATTAAGAAATTTGAGCAAGAGACTGTAAGTCAACAGCAACAGCAAGAA 516
Db 468 GTGAGAACATTTGAGAAAGTCACTAAAGATTTGTCAAGGTTACCAAGGACACATGAG 527
Qy 517 GATTGTAAGCACTTATTAAGCTTATGAGGCTGCTGTTGTAAGGCACTTCTGTAAGCA 576
Db 528 GATGCAAGAAATTTACTTACCTTAATGAGCAATTCCTTATGATGATGATGATGATGAT 587
Qy 577 GAGCAGATGTCAGACCCCTTTGATTAACGATAGGTGCTGCTGCTGCTGCTGCTGCTGCT 636
Db 588 GAGCCAGCTGTCGTCCTTGTAGTAAGACAGCAAGAGTATGCTGACAGCAAGCAAGAC 647
Qy 637 AAGGACTCCCTTACTTTGGGGCTCCACAGTCTCTGCTGATTTAATGATCCAACTTC 696
Db 648 ATGATGCTCTGACCTTTGTCGACCCCTGCTGCTTGTAGGCACTGACAGCAAGCAAG 707
Qy 697 AAGAAATATACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756
Db 708 AAAAAAGCTGCTTATGACAGAGTTCATTAACCTGATGATGATGATGATGATGATGAT 767
Qy 757 ATGACAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 816
Db 768 CATGAAACATTTTGTGATCTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 827
Qy 817 GGTATCGGGGGGCAACAGCTCTGAAACTTATGCTGATGATGATGATGATGATGATGAT 876
Db 828 GGCATTTGAGCCCAAAAGACCATTTGATGATGATGATGATGATGATGATGATGATGAT 887
Qy 877 TTGAGATCTTATTAAGACAGATATCAATTTCTGAGAGTGGCTTCCAGAAAGCT 936
Db 888 ATTTGATACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 947
Qy 937 CGAGCTTGTTCAGAGAGCTTAA--TGTCACTTTGATATTTCTGAGCTAAATAGACT 993
Db 948 AATCATCTTTTCTTGGAAACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1007
Qy 994 GCACCTGATGAGAGGCTCTATTAAGTTCTGCTGTAAGATTAATGTTTCAATGAAGAT 1053
Db 1008 GAGCCGAGGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1067
Qy 1054 CGGGTGACAAAGGCGCATGAGAGATCAATCTGCCAAGATTAATGCTGTCAGAGAGA 1113
Db 1068 CCGATACGCAAGGGGCGCAAAAAGTTGGCCAAAGACCGGCAAGGCAAGGCAAGGCGCG 1127
Qy 1114 CTCGATCTTTTTCAGCCAACTGCCACATCAGCACCCTTAAGACGAGAGACT 1173
Db 1128 CTGGATGATCTTTTTCAGAGTCCGATCCGTTAGCTCACTCAATTAAGAAAGAGCGAGA 1187
Qy 1174 TCGGATTAAGCAAGCAAGCAGCAGCTGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1232
Db 1188 TCCAAAGATCCGCTAAAGAAAGCCAAACTGAGAGTACCCTCGCAGGAAGTTCAA 1246

RESULT 8
AF065397 1357 bp mRNA linear VRT 05-JAN-1999
LOCUS AF065397
DEFINITION Xenopus laevis flap endonuclease 1 (FEN1) mRNA, complete cds.
ACCESSION AF065397
VERSION AF065397.1 GI:4106357

/db_xref="taxon:8355"
 /cell_type="oocyte"
 1. .1149
 /gene="FEN1"
 1. .1149
 /gene="FEN1"
 /function="structure-specific nuclease"
 /product="flap endonuclease 1"
 /protein_id="AB88707.1"
 /db_xref="GI:2674207"
 /translation="MGTHGLAKLADYAPPAIKENDIKSYFGKRVADSMCYOFLI
 AVNDSNTLONEGETTSLHMGFYRTIRVHEGKIPVYVDKPRQMSGELAKRE
 RRAEAKLEAAEEAGEVENIEKTRKLVVTRKOHNEECKLLTLMGIPYVDAPEAE
 ATCAALVKAGKYVAATEEMDALTEGTPVLRHLTASAKRLPJOEFHLNRPVODIGI
 THEOPVDFCLILSDYCEITRGIQPKRAIDILROHNTIDITIDIKKYPENMTH
 KEAKHLEPEVVDITDELKWIPEDEKIVANMCGEKPSRIRIRGAKKLAKNNRG
 STGRLDPEFKVTSYSSSTRKAESEKSKAKKAKTGCTPAGKFKRK"

BASE COUNT 352 a 247 c 306 g 244 t
 ORIGIN

Query Match 20.8%; Score 307; DB 5; Length 1149;
 Best Local Similarity 55.6%; Pred. No. 1.8e-65;
 Matches 633; Conservative 0; Mismatches 500; Indels 6; Gaps 2;

OY 97 ATGGGATCAAGGCTTTGACGAAACTCTGCGGACATATGCGCCCAAGCGCATGAGAG 156
 DB 1 ATGGGAATTCAGCTTTGGCCAAACTTATGCAAGATGTGCACTGCGACCTGACCTATCAAGAG 60
 OY 157 CAGAGTTGAGAGCTCTCTGCGGCAAAATGCGCCCTGACCGCAGCATGAGATCTAC 216
 DB 61 CATGATTTAAAGTTACTTGTGCGAAGGTGGCATGATCGATCTCATGTGCTATTTAT 120
 OY 217 CAGTTCCTGATAGTAGTTGAGAGACAGCATGGAACCTCTCACAATGAGAGCTGTGAA 276
 DB 121 CAGTTCTTATGCGGTGAGACAGAGATGGA--ACACGCTGCAAAAGAAAGATGAA 177
 OY 277 GTGACTAGTATTTGCAAGATGTTCAACCGACATTAAGATTACTGGAAGCGGAAATC 336
 DB 178 ACCACAGGCACTTTATGAGTATGTTTACCGACCATTTGCAATGATGAGATGCGCATC 237
 OY 337 AACGATTTATGTTTGTGCAAGCTCTCTGATATGAAGAAACAAAGATTTGCTAAA 396
 DB 238 AACCAAGTATGTTGTTGATGGAAACCCGACAGATGAAGTCGGTGAATGACCAAA 297
 OY 397 AGTACTCAAAAAGATGATGCAACCAAGATCTGACTAGGCAATGAGATGAGAT 456
 DB 298 CGCAGTAGCGTAGGCGCGAGCAAGAAAGTTGCTTGAAGCTGCGGAGGCGTGGAGAA 357
 OY 457 AAGATGCGATTTGAAAATGAGCAGAGAGACGTGTAAGTTCACAGGCAACACAGAA 516
 DB 358 GTGGAGACATTTAGAAAGTTCACTAAAGACTTGTCAAGGTTACCAAGCAGCAAAATGAG 417
 OY 517 GATTGTAAGCATTATTAAGACTTATGAGGCTCTGTTGTAGAGCACTTGTGAAGCA 576
 DB 418 GAGTGAAGAAATTAATTAACCTAATGGGACTTCTTATGATGATGACCGCTGTGAAGCT 477
 OY 577 GAAGCAATGTGACAGCCCTTTGCATTAACGATTAAGGTGTTGCTGTTGCTGAGAGAT 636
 DB 478 GAGGCCAGGTGCTCTTGAATAAGCAGGGAAGATGATGATCTGACAGCAAGAGAC 537
 OY 637 AAGGATCTCTTCTTTGGGCTCCAGGCTCTCTGCTCATTTAATGATGATCAAGTTC 636
 DB 538 ATGATGCGCTGACCTTTGGACCCCTGTGCTCTTAGGCACTGACAGCAAGCGAAGCT 597
 OY 697 AAGAAAATACCTGTGATGAAATTTGATGTTGCCAAGTTTGGAGAGCTTGAACCTAC 756
 DB 598 AAAAAGCTGCTATTCAGGAGTTCATTAACCGTGTATTAAGGACATAGTATTCACA 657
 OY 757 ATGACCACTTCAATTTGATTTGATGATCCGTGTGATGATGATGATGATGATGATGAT 816
 DB 658 CATGAACAATTTGTGATCTGTGATCTGCTGCAAGGATGATGATGATGATGATGATGAT 717

OY 817 GGTATCGGGGGCAACAGACGTCTGAACTTATGCTCAACATGGGTGCATGAAGCATC 876
 DB 718 GGCATTTGACCCAAAGAGGCACTTACCTGATCCGTGACATTAAGACCATTTGACGATC 777
 OY 877 TTGAGAACTCTTAATTAAGACAGATATCAAAATTCCTGAGACCTGCTTTACCAAGAGCT 936
 DB 778 ATTGATACATTTGACCTCAAGAAATATCCCGTCCGGAACCTGGCTCCACAGAGAGCA 837
 OY 937 CGACCTGTGTTCAAGGAGCCTTAA--TGTGACTTTTGATATTTCTGAGCTAAATGAGCT 993
 DB 838 AAACATCTTTCTTCTTGAACCAAGAGTATGACACCTGACACCTGACAGCTGAAGGTGC 897
 OY 994 GCACCTGATGAGAGGCTCTATAGTTTCTGCTGATTAAGATTAATGTTCAATGAAGAT 1053
 DB 898 GAGGCCATTAAGAAAGATTTGTTGCTTCAATGTTGTTGTTGTTGTTGTTGTTGTTGTT 957
 OY 1054 CGGTCACAAAGCCCATAGAGATCAAAATCTGCCAAGATTAATCTGCCAAGAGAA 1113
 DB 958 CGCATAGCAACGGGGCCAAAGAGTTGGCCAAAGACCGCAAGCAAGCAGCAGAGGCGG 1017
 OY 1114 CTGAGTCTCTTTTCAAGCCCACTGCCACCATCAGACCGCTTAAAGGAGAGACT 1173
 DB 1018 CTGATGACTTTTCAAGATGATGATCGGATCGTTAGCTCACTAAAGAAAGAGCAGAA 1077
 OY 1174 TCGGATAAACCAAGCAGGCTGCGAACAAGAAACAAAGGCTGGTGGAAAGAGA 1232
 DB 1078 TCCAAAGATCCGCTAAAGAAAGCCAAACCTGAGAGTACCCTCGAGGAAGTTCAA 1136

RESULT 10
 BC010203 2163 bp mRNA linear ROD 22-OCT-2001
 LOCUS
 DEFINITION
 Mus musculus, similar to flap structure-specific endonuclease,
 clone MGC:7137 IMAGE:3158176, mRNA, complete cds.
 ACCESSION
 BC010203
 VERSION
 BC010203.1 GI:16307327
 KEYWORDS
 MGC.
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 DIRECT SUBMISSION
 Submitted (05-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobebcm.tmc.edu
 Villalon, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

FEATURES
 source
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plates: 6 Row: d Column: 19.

1. .2163
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:7137 IMAGE:3158176"
 /tissue_type="Mammary tumor, C3(1)-Tag model, infiltrating
 ductal carcinoma, 5 month old virgin mouse."

Quality coverage: agarose-FP - N/A
Quality coverage: 4 x in Q20 bases; sum-of-contigs estimation

NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 34556: contig of 34556 bp in length
* 34557 34576: gap of unknown length
* 34577 62727: contig of 28151 bp in length
* 62728 62747: gap of unknown length
* 62748 81216: contig of 18469 bp in length
* 81217 81236: gap of unknown length
* 81237 92870: contig of 11634 bp in length
* 92871 92890: gap of unknown length
* 92891 103627: contig of 10737 bp in length
* 103628 103648: gap of unknown length
* 103649 114697: contig of 11050 bp in length
* 114718 123385: contig of 8668 bp in length
* 123386 123405: gap of unknown length
* 123406 131815: contig of 8410 bp in length
* 131816 131835: gap of unknown length
* 131836 132037: contig of 202 bp in length
* 132038 132057: gap of unknown length
* 132058 138474: contig of 6416 bp in length
* 138475 144023: gap of unknown length
* 144024 144043: gap of unknown length
* 144044 144924: contig of 881 bp in length
* 144925 145191: contig of 247 bp in length
* 145192 145211: gap of unknown length
* 148414 148434: gap of unknown length
* 148435 151539: contig of 3105 bp in length
* 151540 151559: gap of unknown length
* 151560 154468: contig of 2909 bp in length
* 154469 154488: gap of unknown length
* 154489 156193: contig of 1705 bp in length
* 156194 156213: gap of unknown length
* 156214 157744: contig of 1531 bp in length
* 157745 157764: gap of unknown length
* 157765 159461: contig of 1697 bp in length
* 159462 159481: gap of unknown length
* 159482 160346: contig of 865 bp in length
* 160347 160366: gap of unknown length
* 160367 161824: contig of 1458 bp in length
* 161825 161844: gap of unknown length
* 161845 163663: contig of 1819 bp in length
* 163664 163684: gap of unknown length
* 163685 164031: contig of 348 bp in length
* 164032 164051: gap of unknown length
* 164052 164700: contig of 649 bp in length
* 164701 164721: gap of unknown length
* 164722 166088: contig of 1368 bp in length.
```

FEATURES

source

1. 166088

Location/Qualifiers

/organism="Mus musculus"

/strain="129/SVETACBr"

/db_xref="taxon:10090"

/chromosome="19"

/clone="RP22-325J22"

/sex="male"

1. 34556

/note="assembly_name:Contig81"

34577..62727

/note="assembly_name:Contig80"

62748..81216

misc_feature

misc_feature

misc_feature

/note="assembly_name:Contig79"
81237..92870

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

BASE COUNT 44602 a 38872 c 39481 g 42573 t 560 others
ORIGIN

Query Match 20.6%; Score 303.8; DB 2; Length 166088;
Best Local Similarity 56.2%; Pred. No. 2.8e-64;
Matches 641; Conservative 0; Mismatches 482; Indels 18; Gaps 3;

```
QY 97 ATGGGCAATGCAAGGGTTTGACGAACCTGCTGGCGGACATGCGCCCAAGGCATGAGAG 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11414 ATGGGAATTCACGGCTTGGCCAAACTAATTCATGATGCGCCCAAGGCATGAGAG 11473
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 157 CAGAGTTGCGAGGACCTACTTGGCCGCAAAATCCCGTGCAGGCGACGATGAGATCTAC 216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11474 AATGACATCAAGACACTTCTTGGTGCAGAACTGCGCATCGCTCTCATATGATCTAC 11533
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 217 CAGTCTGATAGTAGTTGGAAGGACGATGGAAGAACTCTACAAATGAAGCTGTGAA 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11534 CAGTCTGATAGTAGTTGGAAGGATGTCACCGGACATGATGAGAAATGAGATC 11590
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 GTCACTAGTCATTTGCAAGGATGTCACCGGACATGATGAGAAATGAGATC 336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11591 ACCACCGACCACTGATGGGCGATGTTTACCGTACCATCGCATGATGAGAAATGAGATC 11650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 AAGCCAGTTATGTTTGTGATGCGACGCTCTGATATGAAGAACAAGAACTTGTAA 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11651 AAGCCTGTAGCTCTTGTGATGCGCAACACGACAGCTGAAGTCAAGGAGCTGCCAAG 11710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 397 AGATCTCAAAAAGATGATGACACCAAGATCTGACTGAGGACGATGAGTGAAGAT 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Db 11711 CGCAGTATGAGGCGCCGCGGCTGAGAAAGCACTGACAGAGGCTTACAGAGGCTGGGATG 11770
QY 457 AAGATGCGATTGAAAAATTGAGCAAGAGAGCTGTAAAGTTCACAAAGCAACAACGAA 516
Db 11771 GAGAGAGAGGTGAGAAAGTTCACCAAGAGGCTGTGAGGTACCAAGAACACAAATGAT 11830
QY 517 GATTGTAACGACTATTAACTTATGAGGGGTTCTGTTTGAAGCACCCTTCTGAAGCA 576
Db 11831 GAGTGCAAACACCTGCTGAGACCTTCATGGGATCCCTTACCTTATGATGACCCAGGAGCA 11890
QY 577 GAACAGATGATGAGCGCTTTGATTAACGATAGGTTCCTGCTGCTGATGAGAAAT 636
Db 11891 GAGGCGACGCTGTGCTGCTGCGCAAGAGGCTGCGAAATCTATCTGCGCGCAGAGAGAC 11950
QY 637 AAGGACTCCCTTACTTTTGGGGCTCCACGGTCTCTGCTGATTTAATGATCCAGTTCC 696
Db 11951 ATGAGACTGCTCTACTTTTGGACAGCCCGCTGCTAATGCGACATTAAGTCCAGTGAAGCC 12010
QY 697 AAGAAATACCTGTGATGGAATTTGATGTCAGAGTTTGAAGAGAGCTTGAACATCACC 756
Db 12011 AAGAACTGCGCCATCCAGAGTTCCATGAGCCGCTCTGCAAGAGAGCTGGTCTGAGC 12070
QY 757 ATGAGACAGTTCATGATTTGATGCTGCTGATGATGATGATGATGATGATGATGATGAT 816
Db 12071 CAGAGAGAGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12130
QY 817 GGTATCGGGGGGCAACAGCTCTGAAATCTTATTCGTCACATGAGGCTCAGTGAAGCAATC 876
Db 12131 GGCATTTGGGGCCCAAGCGGCTGTGATCTCATCCAGAAACATTAAGACATCGAGGAGATC 12190
QY 877 TTGAGATCTTTAATAAGACAGATATCAATTCCTGAGGAGCGGCTTACCAAGAAAGCT 936
Db 12191 GTGAGGCGGCTGGAGCCCAAGAGTACCCGCTTCAGAGAGTGTGCTCCAGAGAGAGCC 12250
QY 937 CGAGCTTTGTCAGAGAGCCCTAATG---TCACCTTTGATTTCTGAGCTAATTAATGAGCT 993
Db 12251 CAGAGAGCTTCTCTGAGAGCCAGAAAGTACTGAGCCAGAGTCTGTGAGAGTGAAGTGAAG 12310
QY 994 GCACCTATGAGAGGAGGCTCATTAAGTTTCTGTTAAATATATGTTTCAATGAAGAT 1053
Db 12311 GAGCCAAATGAAAGAGAGTTGTCATTAATTTATGTTGTTAAAGCAAGTTTCTGTAAGAG 12370
QY 1054 CGGCTGCAAAAGGCGCATAGAGAGATCAATTCGCCAAGATTAATTCGTCGAAGAGAGA 1113
Db 12371 CGATTGCGAGTGGGGTCAAGCGGCTGATGAAGAGCGCGCAGAGGAGAGAGAGAGAGC 12430
QY 1114 CTGAGTCTTTTTCAGAGCCAACTGCGACCATCATGACGCGCTAATAACGGAAGAGACT 1173
Db 12431 CTCGATATTTCTTCAAGGTGACAGGCTCACTCTCTCAGC-----TAAGCGC 12478
QY 1174 TCGATTAATAACAGCAGAGCGAGTGCAGAACAAAGAAAGAGCTGTGTGAAGAAAGAA 12538
Db 12479 AAGGAGCCAGAACCCAGAGGCGCTGTTAAGAGAAAGCAAGAGTGGGAGCGGAGAG 12538
QY 1234 T 1234
Db 12539 T 12539

```

```

RESULT 12
HSRAD2      1757 bp  mRNA  linear  PRI 07-DEC-1998
LOCUS      Homo sapiens mRNA for flap endonuclease-1.
DEFINITION X76771.1 GI:1905802
VERSION    X76771.1
KEYWORDS   feni gene; flap endonuclease-1.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota: Chordata: Craniata: Vertebrata: Euteleostomi:
            Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo.
REFERENCE  1 (bases 1 to 1757)
AUTHORS   Watts,F.

```

```

TITLE      Direct Submission
JOURNAL    Submitted (16-DEC-1993) F. Watts, University of Sussex, School of
            Biological Sciences, Falmer, Brighton BN1 9QG, UK
REFERENCE  2 (bases 1 to 1757)
AUTHORS    Murray,J.M., Tavassoli,M., al-Harithy,R., Sheldrick,K.S.,
            Lehmann,A.R., Carr,A.M. and Watts,F.Z.
TITLE      Structural and functional conservation of the human homolog of the
            Schizosaccharomyces pombe rad2 gene, which is required for
            chromosome segregation and recovery from DNA damage
JOURNAL    Mol. Cell. Biol. 14 (7), 4878-4888 (1994)
MEDLINE    94277093
FEATURES   location/Qualifiers
            source      1..1757
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        368..1510
                        /gene="feni1"
                        368..1510
                        /gene="feni1"
                        /codon_start=1
                        /product="flap endonuclease-1"
                        /protein_id="CA54166.1"
                        /db_xref="GI:3980293"
                        /translation="MGIOGLAKLIADYAPSAIRENDIKSYFGRKVAIDASMSIYQPLI
                        AVROGQVLOMEGETSHLMGMYRTIRRMENGIRKPYVFDKRPOLKSGELAKRSE
                        RRAERKLOIOAAGAEOMEYKFKTRLVTVTOHDECKHLISLIPYLDAPSAE
                        ASCALVRAKQKYAAATREMDCLTFGSPVLRHLTASEAKLPIDRPHLSRIIOELG
                        NOEOPVDICILISDYCESIRGIAPKRAVDLIQKHSIEIVRLDPNKYPVENNLH
                        KEAHLDFLEPVLDPSEVLEKMSPEENELIKFMCGKQSEIRISGVARKLSKSHOG
                        STQRLDDFFVYVGLSLSARKKEDEPFGSTTKRAKATGAAGKFRGK"
BASE COUNT 418 a 443 c 542 g 354 t
ORIGIN
Query Match 20.3%. Score 299.6; DB: 9; Length 1757;
Best Local Similarity 54.8%; Pred. No. 1,4e-63;
Matches 659; Conservative 0; Mismatches 354; Indels 9; Gaps 3;
QY 97 ATGGCATCAAGGGTTTACGAAACTGCTGGCGGCAATGCGCCCAAGGCGCATGAAGAG 156
Db 368 ATGGCAATTCAGAGGCTGCGCAACTAATTTGCTGATGTGGCCGCCAGTCCGCGAG 427
QY 157 CAGAAATTCGAGAGCTACTTGGCGCGCAAAATTCGCGCTGACGCCAGCATGATCATC 216
Db 428 AATGACATCAAGAGCTACTTGGCCGTAAAGGTGGCCATGATGCTTATGAGCATTTAT 487
QY 217 CAGTCTCGATAGTGTGGAAGGACGATGGAACATCTCACAAATGAAGTGTGGA 276
Db 488 CAGTCTCGATAGTGTGGAAGGACGATGGAACATCTCACAAATGAAGTGTGGA 544
QY 277 GTCACTAGTCATTTCAAGGATTTCAACGAGGACGATTAAGATTACTGGAAGCGGAGATC 336
Db 545 ACCACGAGCAGCTATGAGGATGTTTACGACACCATTTGCGATGATGAGAGAGGATC 604
QY 337 AAGCCAGTTTATGTTTGTGATGCAAGCTCTCTGATTAAGAAAGAAAGAACTTGCTAAA 396
Db 605 AAGCCGTTGATGTTTGTGATGCAAGCGCCACAGCTCAAGTCAAGTCAAGGAGCTGCGCAA 664
QY 397 AGATACTCAAAAAGAGATGATGACCAACAAAGATCTGATGAGCGAGTGAAGTGAAGAT 456
Db 665 CCGAGTGAAGCGGCGGCTGAGGACAGAGACAGCTGCGACGAGGCTCAGGCTGCTGGGCC 724
QY 457 AAGATGCGATTGAAAAATTGAGCAAGAGAGCTGTAAAGTTCACAAAGCAACAACGAA 516
Db 725 GAGCAGAGAGGTGAGAAATTTCACTAAGCGGCTGTGGAAGTCACTAAGCAGACAAATGAT 784
QY 517 GATTGTAACGACTATTAACTTATGAGGGGTTCTGTTTGAAGCACCCTTCTGAAGCA 576
Db 785 GAGTGCAAACACCTGCTGAGAGCTCATGGGCAATCCCTTATTTGAAGCACCAGAGAGCA 844
QY 577 GAGCAGATGATGAGCGCTTTGATTAACGATAGAGTGTGCGTGTGCTTCAAGAGAT 636
Db 845 GAGCGAGCTGTGCTGCGCTGTGTAAGGCTGCGCAAGTCTATGCTGCGGCTAACGAGAGAC 904

```


* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 20726: contig of 20726 bp in length
* 20727 20826: gap of 100 bp
* 20827 34666: contig of 13840 bp in length
* 34667 34766: gap of 100 bp
* 34767 46065: contig of 11299 bp in length
* 46066 46165: gap of 100 bp
* 46166 53516: contig of 7351 bp in length
* 53517 53616: gap of 100 bp
* 53617 61065: contig of 7449 bp in length
* 61066 61165: gap of 100 bp
* 61166 67998: contig of 6833 bp in length
* 67999 68098: gap of 100 bp
* 68099 74526: contig of 6428 bp in length
* 74527 74626: gap of 100 bp
* 74627 81299: contig of 6673 bp in length
* 81300 81399: gap of 100 bp
* 81400 87316: contig of 5917 bp in length
* 87317 87416: gap of 100 bp
* 87417 91541: contig of 4125 bp in length
* 91542 91641: gap of 100 bp
* 91642 95733: contig of 4090 bp in length
* 95732 95831: gap of 100 bp
* 95832 101128: contig of 5297 bp in length
* 101129 101228: gap of 100 bp
* 101229 104948: contig of 3720 bp in length
* 104949 105048: gap of 100 bp
* 105049 110306: contig of 5258 bp in length
* 110307 110406: gap of 100 bp
* 110407 112969: contig of 2563 bp in length
* 112970 113069: gap of 100 bp
* 113070 116484: contig of 3415 bp in length
* 116485 116584: gap of 100 bp
* 116585 119599: contig of 3015 bp in length
* 119600 119699: gap of 100 bp
* 119700 121881: contig of 2182 bp in length
* 121882 121981: gap of 100 bp
* 121982 122703: contig of 722 bp in length
* 122704 122803: gap of 100 bp
* 122804 125566: contig of 2763 bp in length
* 125567 125666: gap of 100 bp
* 125667 126845: contig of 1179 bp in length
* 126846 126945: gap of 100 bp
* 126946 129349: contig of 2404 bp in length
* 129350 129449: gap of 100 bp
* 129450 131268: contig of 1819 bp in length
* 131269 131368: gap of 100 bp
* 131369 132565: contig of 1197 bp in length
* 132566 132665: gap of 100 bp
* 132666 134031: contig of 1366 bp in length
* 134032 134131: gap of 100 bp
* 134132 136176: contig of 2045 bp in length
* 136177 136276: gap of 100 bp
* 136277 137418: contig of 1142 bp in length
* 137419 137518: gap of 100 bp
* 137519 138627: contig of 1109 bp in length.
```

FEATURES

source

```
1.138627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="CMB9-22K5"
1.20726
/misc_feature
/note="assembly_fragment"
20827.34666
/misc_feature
/note="assembly_fragment"
34767.46065
/misc_feature
/note="assembly_fragment"
```

```
misc_feature 46166..53516
/note="assembly_fragment"
misc_feature 53617..61065
/note="assembly_fragment"
misc_feature 61166..67998
/note="assembly_fragment"
misc_feature 68099..74526
/note="assembly_fragment"
misc_feature 74627..81299
/note="assembly_fragment"
misc_feature 81400..87316
/note="assembly_fragment"
misc_feature 87417..91541
/note="assembly_fragment"
misc_feature 91642..95731
/note="assembly_fragment"
misc_feature 95832..101128
/note="assembly_fragment"
misc_feature 101229..104948
/note="assembly_fragment"
misc_feature 105049..110306
/note="assembly_fragment"
misc_feature 110407..112969
/note="assembly_fragment"
misc_feature 113070..116484
/note="assembly_fragment"
misc_feature 116585..119599
/note="assembly_fragment"
misc_feature 119700..121881
/note="assembly_fragment"
misc_feature 121982..122703
/note="assembly_fragment"
misc_feature 122804..125566
/note="assembly_fragment"
misc_feature 125667..126845
/note="assembly_fragment"
misc_feature 126946..129349
/note="assembly_fragment"
misc_feature 129450..131268
/note="assembly_fragment"
misc_feature 131369..132565
/note="assembly_fragment"
misc_feature 132666..134031
/note="assembly_fragment"
```

Query Match 20.3%: Score 299.6; DB 2; Length 138627;
Best Local Similarity 54.8%; Pred. No. 3e-63;
Matches 659; Conservative 0; Mismatches 534; Indels 9; Gaps 3;

```
QY 97 ATGGGCATCAAGCGTTTGACGAACCTGCTGGCGGACAAATGCGCCGACAGCCGACGATGAGGAG 156
    ||||| || || || || || || || || || || || || || || || || || || || || || ||
Db 17322 ATGGCAATTCGAAGGCGCTGCCAACTAATTCCTGATGTGGCCCCCGACATCGCGGAG 17381

QY 157 CAGAAGTTCGAGAGCTATTGGCGCGCAAAATCCGCGTCGACGCGCAGATGACATCTAC 216
    || || || || || || || || || || || || || || || || || || || || || || || ||
Db 17382 AATGNCATCAAGAGCTACTTTGGCCGTAAAGTGGCCATGTGATGCTTATGAGCATTTAT 17441

QY 217 CAGTTCCTGATAGTATGTGAAGGACAGCATGCAAACTCTCACAAATGAAGCTGTGTA 276
    ||||| ||||| || || || || || || || || || || || || || || || || || || || || ||
Db 17442 CAGTTCCTGATGTGCTGCCGAC---GGTGGGATGTGCTGCGCAAAATGAGAGGGGTAG 17498

QY 277 GTCACTAGCTCAATTTGCAAGCATGTTCAACCGGACATAGATTACTGGAAGCGGAATC 336
    || || || || || || || || || || || || || || || || || || || || || || || ||
Db 17499 ACCACAGCCACCTGATGGGCGATGTTACCGCACCATTCATGATGAGAGACGGCATC 17558

QY 337 AAGCCAGTTTATGTTTGTGATGGCAAGCGCTCGTATGTAAGAACAACAACTGCTTAA 396
    ||||| || ||||| ||||| ||||| || || || || || || || || || || || || || || ||
Db 17559 AAGCCCGTATGTCTTTGATGAGCAAGCCGACAGCTCAAGTCAAGCGGAGCTGGCCAAA 17618

QY 397 AGATAGTCAAAAAGATGATGCAACCAAGATCTGACTGAGGAGATGAGATGAGAT 456
    | || || || || || || || || || || || || || || || || || || || || || || ||
Db 17619 CGCAGTAGCGGCGGCTGAGAGCAGAGAGAGCTGCAGCAGGCTCAAGGCTGCGGGGCC 17678

QY 457 AAAGATGCGATTGAATAATTGAGCAAGAGCACTTAAAGGTCACAGGCAACACACGAA 516
```

```

Db 17679 GAGCAGAGAGTGGAAAAATTCACCTAAGCGCTGGTGAAGTCTCTAAGCAGACATATGAT 17738
Qy 517 GATTGTAAACGACTATTAACTTAAAGGGGGTCTCTGTTAGAGGCACTTCTGAAGCA 576
Db 17739 GAGTGGAAACATCTGCGTGCAGCCCTCATAGGGATCCCTATCTTGATGACCCAGTAGGCA 17798
Qy 577 GAAGCAGATATGTCAGCCCTTTGCTATTAACGATAGGTGTCCCTGTCTTCAAGAT 636
Db 17799 GAGGCCAGTGTGCTGCTGCGCTGGTGAAGGCTGGCAAGTCTATCTCGCTACCGAGGAC 17858
Qy 637 AAGGACCTCCCTTACTTTTGGGCTCCAGCGTCTCTGCTGCTATTAATGATCCAACTTCC 696
Db 17859 ATGACGACCTCTCAGCTTCCGAGCCCTGTCTCTATGCGACACTGACTGCCATGAAAGCC 17918
Qy 697 AAGAAATACCTGTGATGATTAATTTGATGTTGCCAAGTTTGGAGAGCTTGAATCCACC 756
Db 17919 AAAAAGCTGCCAATCCAGGAATTCACCTGAGCCGGATTCCTGAGAGAGCTGGGCTGAAAC 17978
Qy 757 ATGACACAGTTCAATGATTTGTGATCCCTGATGAGTGTGATGATGATGATGATGATGAT 816
Db 17979 CAGGAAACAGTTTGTGATCTGTGATCTCTCTGAGCAGTACTACTGAGTATATCCGG 18038
Qy 817 GGTATCGGGGGGCAAAAGCTCTGAACCTTAATTCGTCAACATGGTCCATAGAAAGCATC 876
Db 18039 GGTATTTGGGCCCAAGGGGGCTGTGACCTCATCAGAAAGCAAGACATCGAGAGATC 18098
Qy 877 TTGGAAGATCTTAAATTAAGACAGATATCAAAATTCCTGAGAGCTGGCTTACCAAGACT 936
Db 18099 GTGGGGGAGTTGACCCCAAGTACCTGCTGCGCAAAATTTGGCTCCCAAGGAGGCT 18158
Qy 937 CGAGCTTTGTCAGAGAGCCCTAATG---TCACCTTGGATATTCCTGAGCTAAATAGGACT 993
Db 18159 CACCAGCTCTTCTTGAACCTGAGTCTGAGCCAGAGTCTGTGAGCTGGAAGTGAGAGC 18218
Qy 994 GCACCTGATGAGAGGCTCTCATAGTTTCTGCTAAAGATTAATGTTTCAATGAAGAT 1053
Db 18219 GAGCCAATATGAAGAAGAGCTGATCAAGTCTGATGTTGTTGTTGTTGTTGTTGTTGTTG 18278
Qy 1054 CGGGTGCACAAAGGCTATAGAACATCAATCTGCCCAAGAAATTAATGCTGCCAAGAGAGA 1113
Db 18279 CGAATCCGCACTGGGGTCAAGAGGCTGATGAGAGCCGCCAAGGACACCCAGGGCCGC 18338
Qy 1114 CTGCAAGTCTTTTTCAGAGCCAACTGCCAC---CACATCAGACCCGCTTAAACGGAAGAG 1170
Db 18339 CTGATGATTTCTTCAAGGTGACCGGCTCACTCTTCACTAGTAAAGGAGAGCCGAA 18398
Qy 1171 ACTTCGATTAACACAGCAGGCTGCGACACAGAAACAAAGGCTGGTGAAGACAG 1230
Db 18399 CCCAAGGATCCACTAAGAAAGAGCAAGACTGGGACACAGGAGTTTAAAGGGGA 18458
Qy 1231 AAATATCTTGGATGCTTATGATGATCACTACGATACGAAAGAGCGGTGATGATCAGC 1290
Db 18459 AAATTAATGTGTTCCCATATATACCTCTTCAACCCAGAAATATTTGCCGTTGTATCC 18518
Qy 1291 TT 1292
Db 18519 TT 18520

```

```

RESULT 15
AP002380/c 169053 bp DNA linear HTG 31-MAY-2000
LOCUS Homo sapiens chromosome 11 clone RP11-467L20 map 11q, WORKING DRAFT
DEFINITION SEQUENCE, 35 unordered pieces.
ACCESSION AP002380.1 GI:8131644
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens DNA, clone:RP11-467L20.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 169053)

```

```

AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Matsunabe,H. and Sakaki,Y.
TITLE
Homo sapiens 169,053 genomic DNA of 11q
JOURNAL
Published Only in Database (2000) In press
REFERENCE
2 (bases 1 to 169053)
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Matsunabe,H. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8535,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
COMMENT
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humdrfl1
Center clone name: RP11-467L20
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 147827 bases at least Q40
Consensus quality: 157936 bases at least Q30
Consensus quality: 162410 bases at least Q20
Insert size: 165653; sum-of-ctrls
Quality coverage: 4.43x in Q20 bases; sum-of-ctrls
-----
NOTE: This is a 'working draft' sequence. It currently consists of
35 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 21322 contig of 21322 bp in length
21423 18825 contig of 18825 bp in length
40346 50492 contig of 50492 bp in length
50593 58569 contig of 58569 bp in length
58670 67385 contig of 67385 bp in length
67486 75248 contig of 75248 bp in length
75349 84453 contig of 84453 bp in length
84554 89492 contig of 89492 bp in length
89593 96445 contig of 96445 bp in length
96546 102631 contig of 102631 bp in length
102732 107199 contig of 107199 bp in length
107300 111820 contig of 111820 bp in length
111921 117068 contig of 117068 bp in length
117169 120901 contig of 120901 bp in length
121002 124072 contig of 124072 bp in length
124173 127611 contig of 127611 bp in length
127712 129401 contig of 129401 bp in length
129502 133141 contig of 133141 bp in length
133242 136380 contig of 136380 bp in length
136481 139425 contig of 139425 bp in length
139526 143281 contig of 143281 bp in length
143382 146449 contig of 146449 bp in length
146550 148828 contig of 148828 bp in length
148929 151045 contig of 151045 bp in length
151146 153372 contig of 153372 bp in length
151473 154996 contig of 154996 bp in length
155097 156763 contig of 156763 bp in length
156864 159133 contig of 159133 bp in length
159233 161349 contig of 161349 bp in length
161450 162850 contig of 162850 bp in length
162951 164533 contig of 164533 bp in length
164634 165829 contig of 165829 bp in length
165930 167130 contig of 167130 bp in length
167231 168290 contig of 168290 bp in length

```


Db 3399 GAGCAGAGAGCTGCAAAATTCACCTAAGCCGCTGTGAGGTCTACCTAAGCAGCACAATGAT 3340
QY 517 GATTGTAAAGACTATTAAACCTTAAGGGGTTCTGTTTAAAGGACCTTCTGAAGCA 576
Db 3339 GAGTGAACAACATCTGCTGAGCCTCATAGGCATCCCTTATCTGTATGACACCAGTGAAGCA 3280
QY 577 GAAGCAGAAATGTGACGCCCTTTGCTAATAGATAGGTTCGCTTGCCTTCAGAAAGAT 636
Db 3279 GAGGCCAGCTGTGTGCTGCTGCTGGAAGGCTGGCAAGTCTATGCTGCGGTACCGAGGAC 3220
QY 637 AAGGACTCCCTTACTTTTGGGGCTCCAGGCTTCCTTCATTATTAATGATCCAGTCC 696
Db 3219 ATGAGCTGCTCACTTCGGCAGGCCCTGTGCTAATGCGACACCTGACTGCCAGTGAAGCC 3160
QY 697 AAGAAATFACCTGTGATGAAATTTGATGTGCCAAGGTTTGGAGAGCTTGAATCAACC 756
Db 3159 AAAAGCTGCCAATCCAGCAATTCACCTGAGCCGGATTCGTGAGGAGCTGGGCTGAAC 3100
QY 757 ATGAGCAGATTTCATTTGATTTGTGATCCTGTGTGATGTGATGATTTGATATCAACAA 816
Db 3099 CAGGAAACAGTTTGTGATGTGTGATCTCTGTGAGCGCATCTACTAGTGTGAGATATCCGG 3040
QY 817 GGTATCGGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGTCCATAGAAAGCATC 876
Db 3039 GGTATTTGGGCCCAACCGGGCTGTGACCTCATCCAGAAAGCACAAGCATCGAGGATTC 2980
QY 877 TTGGAATATCTTATTAAGACAGATATCAATTCCTGAGAGCTGGCCTTACCAAGAACT 936
Db 2979 GTGCGGGGACTTTGACCCCAACAGTACCTGTGTGCCAGAAAATTTGGCTCCACAAGAGGCT 2920
QY 937 CGAGCCTTGTTCAGAGAGCCTAATG--TCACTTTGGATATTCCTGAGCTAATATGACT 993
Db 2919 CAGCAGCTCTTCTGGAACCTGAGTGTCTGAGCCACAGTCTGTGAGCTGGAAGTGGAGC 2860
QY 994 GCACCTGATGAGAGGCTCTCATTAAGTTCTGTGTAAGATTAATGTTTCAATGAAGAT 1053
Db 2859 GAGCCAATGAAGAGAGCTGATCAAGTTGATGTGTGAAGAGCAGTTCTCTGAGGAG 2800
QY 1054 CGGGTGACAAAGGCCATAGAGAAATCAAAATCTGCAAGAAATTAATGTCGAAGGAGA 1113
Db 2799 CGAATCCGCACTGGGCTCAAGAGGCTGATAGAGCCGCCAAGGCAAGCACCAGGCGCC 2740
QY 1114 CTCGAGTCTTTTTCAGGCCAATGCCAC--CACATCAGCACCGCTTAAACGGAAGAG 1170
Db 2739 CTGGATGATTTCTTCAAGGTACCGGCTCACTCTTTCAGCTAAGCCGAGAGCCAGAA 2680
QY 1171 ACTTCGATTAACACAGAGGAGCTGCGAACAAGAAACAAAGGCTGCTGGAAGAAG 1230
Db 2679 CCGAAGGGATCCACTAAGAAAGAGCAAGCTGGGCGACGAGGAGGATTAAAGGGGA 2620
QY 1231 AAATATCTTGGATGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1290
Db 2619 AAATTAATGTGTTTCCCATATATACCTCTTCAACCCAGAAATATTTGCCGTTGTACCC 2560
QY 1291 TT 1292
Db 2559 TT 2558

Search completed: November 5, 2002, 13:04:46
Job time : 2474.66 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 08:10:48 : Search time 176.457 Seconds
(without alignments)
14380.822 Million cell updates/sec

Title: US-09-805-311-7

Perfect score: 1478
Sequence: 1 cgaccacgcgtccgggaaa.....aaaaaaaaaaaaaaaa 1478

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_032802:*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1478	100.0	1478	21	AAAZ27926
2	1421.8	96.2	1463	21	AAAZ27923
3	1401.2	94.8	1541	21	AAAZ27924
4	1334.2	90.3	1381	21	AAAZ27925
5	299.4	20.3	1144	20	AAAZ2117
6	299.4	20.3	1144	20	AAAZ2107
7	275	18.6	1300	23	AAAZ02111
8	274.4	18.6	1300	23	ABLI14287
9	270.6	18.3	1330	20	AAAZ02108

10	265.4	18.0	1149	20	AAAZ02109
11	251	17.0	3362	23	ABLI14286
12	173.2	11.7	5471	23	ABLI10094
13	162.8	11.0	1023	18	AAAZ76685
14	162.8	11.0	1023	19	AAAZ5840
15	162.8	11.0	1023	19	AAAZ5951
16	154.4	10.4	1054	19	AAAZ5984
17	152.8	10.3	1115	19	AAAZ5992
18	138.8	9.4	34980	22	AAAZ1224
19	137.2	9.3	1032	20	AAAZ1850
20	136.4	9.2	510	22	AAAZ37479
21	136.4	9.2	510	22	AAAZ18286
22	130.4	8.8	1115	19	AAAZ5978
23	129.6	8.8	514	19	AAAZ5990
24	128	8.7	1164	19	AAAZ5975
25	114.4	7.7	1729	23	AAAZ83732
26	111.8	7.6	889	19	AAAZ5974
27	100.8	6.8	1164	19	AAAZ5989
28	98.6	6.7	1011	19	AAAZ5913
29	98.6	6.7	1011	19	AAAZ54009
30	94.8	6.4	386	19	AAAZ5979
31	92.4	6.3	981	18	AAAZ76682
32	92.4	6.3	981	19	AAAZ5948
33	90	6.1	546	22	AAAZ3065
34	90	6.1	546	22	AAAZ11490
35	90	6.1	546	22	AAAZ37269
36	90	6.1	546	22	AAAZ18098
37	89.2	6.0	981	19	AAAZ5837
38	85	5.8	245	22	AAAZ50354
39	85	5.8	245	22	AAAZ12416
40	79.4	5.4	777	19	AAAZ54010
41	77.8	5.3	840	19	AAAZ5977
42	72.2	4.9	987	19	AAAZ54011
43	69.4	4.7	296	19	AAAZ5976
44	64.4	4.4	514	19	AAAZ5985
45	64	4.3	4551	21	AAAZ70129

ALIGNMENTS

RESULT 1	AAAZ27926	standard: cDNA; 1478 BP.
ID	AAAZ27926	
AC	AAAZ27926:	
XX		
DT	12-SEP-2000	(first entry)
XX		
DE	Maize Rad2/FEN-1 CDNA.	
XX		
KW	Maize: Rad2/FEN-1; transgenic plant; male sterile plant;	
KW	endonuclease; exonuclease; DNA repair; gene targeting; ss.	
XX		
OS	Zea mays.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	97..1236
XX		/tag= a
XX		
PN	WO200036109-A1.	
XX		
PD	22-JUN-2000.	
XX		
PF	16-NOV-1999;	99WO-US27147.
XX		
PR	15-DEC-1998;	98US-0112332.
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Mahajan PB;	
XX		
DR	WPI; 2000-452026/39.	

Yeast FEN-1 CDNA.
Drosophila melanog
Drosophila melanog
Pyrococcus furiosu
Pyrococcus furiosu
Nucleotide sequenc
P. furiosus N-term
Pyrococcus N-term
Pyrococcus heat re
Human bone marrow
Probe #8219 for ge
P. furiosus N-term
Nucleotide sequenc
Nucleotide sequenc
DNA encoding novel
Nucleotide sequenc
M. jannaschii N-te
Archaeoglobus fulg
Nucleotide sequenc
Nucleotide sequenc
Methanococcus jann
Nucleotide sequenc
Human foetal liver
Human brain expres
Human bone marrow
Probe #8031 for ge
Methanococcus jann
Human bone marrow
Probe #17349 for g
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Nucleotide sequenc
Plasmodium talcipa

DR P-PSDB: AAY95310.

XX Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -

XX Example 1; Page 77-79; 85bp; English.

XX The present sequence is that of maize cDNA coding for Rad2/FEN-1
CC (see AAY95310). The cDNA was isolated from a library prepared
CC from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a
CC structure specific endonuclease which under certain conditions also
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.

XX Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;

Query Match 100.0%; Score 1478; DB 21; Length 1478;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCCAGCGCTCGGGAATAGCTCGCGGTTCCTGCGCAGCTCGCGCTCAGC 60
DB 1 CGACCCAGCGCTCGGGAATAGCTCGCGGTTCCTGCGCAGCTCGCGCTCAGC 60
QY 61 CGCGCCCGCCAGCCGCGCAGCCGCGCAGAGAGATGGGATCAAGGGTTGACGAA 120
DB 61 CGCGCCCGCCAGCCGCGCAGCCGCGCAGAGATGGGATCAAGGGTTGACGAA 120
QY 121 CTGCTGGGGGACATGGGCGCAAGCGATGAGAGAGATGCAAGCTACTCGGC 180
DB 121 CTGCTGGGGGACATGGGCGCAAGCGATGAGAGAGATGCAAGCTACTCGGC 180
QY 121 CTGCTGGGGGACATGGGCGCAAGCGATGAGAGAGATGCAAGCTACTCGGC 180
DB 121 CTGCTGGGGGACATGGGCGCAAGCGATGAGAGAGATGCAAGCTACTCGGC 180
QY 181 CGCAAAATCGCGCTCGAGCGCGCATGAGCATCTACCACTTCTGATAGTTGAAGG 240
DB 181 CGCAAAATCGCGCTCGAGCGCGCATGAGCATCTACCACTTCTGATAGTTGAAGG 240
QY 241 ACAGGATGGAATCTTCACAAATGAAGCTGGAAGTCTAGTATTTGGCAAGAAATG 300
DB 241 ACAGGATGGAATCTTCACAAATGAAGCTGGAAGTCTAGTATTTGGCAAGAAATG 300
QY 301 TTCAACCGGAAATAGATTACTGGAAGCGGAAATCAAGCAGTTATGTTTTGATGGC 360
DB 301 TTCAACCGGAAATAGATTACTGGAAGCGGAAATCAAGCAGTTATGTTTTGATGGC 360
QY 361 AAGCTCTGATATGAGAAACAGAACTTGCTTAAAGATACCAAAAAGATGATGCA 420
DB 361 AAGCTCTGATATGAGAAACAGAACTTGCTTAAAGATACCAAAAAGATGATGCA 420
QY 421 ACCAAAGATCTGAGTGAAGGAGTAGAGATGAAGATCGATTGAAAATGAGC 480
DB 421 ACCAAAGATCTGAGTGAAGGAGTAGAGATGAAGATCGATTGAAAATGAGC 480
QY 481 AAGAGAGCTGTTAAAGGTCAAGGCAACACACAGAGATTGTAAAGCATTTAAGACTT 540
DB 481 AAGAGAGCTGTTAAAGGTCAAGGCAACACACAGAGATTGTAAAGCATTTAAGACTT 540
QY 541 ATGGGGGTTCTGTTGTAGAGACACCTTTGAAGCAGACAGATGTGAGCCCTTTGG 600
DB 541 ATGGGGGTTCTGTTGTAGAGACACCTTTGAAGCAGACAGATGTGAGCCCTTTGG 600

QY 601 ATAAAGATPAAGAGTGTTCCTGTTGCTGAGAGATPAAGAGCTCCCTTTGGGGCT 660
DB 601 ATAAAGATPAAGAGTGTTCCTGTTGCTGAGAGATPAAGAGCTCCCTTTGGGGCT 660
QY 661 CCACGGTTCCTGCTCAATTAATGATCCAAGTCCACAGAAATACCTGTGATGCAATTT 720
DB 661 CCACGGTTCCTGCTCAATTAATGATCCAAGTCCACAGAAATACCTGTGATGCAATTT 720
QY 721 GATGTGCGCAAGGTTTGGAGGAGCTTGAACCTGACCATGAGCAGTTCATGATTTGTC 780
DB 721 GATGTGCGCAAGGTTTGGAGGAGCTTGAACCTGACCATGAGCAGTTCATGATTTGTC 780
QY 781 ATCTGTGTGATGTGATATTTGATAGCATCAAGATATCGGGGGCAACAGCTCTG 840
DB 781 ATCTGTGTGATGTGATATTTGATAGCATCAAGATATCGGGGGCAACAGCTCTG 840
QY 841 AAATCTTGTGCAACATGAGTGGCTCCATGAGAAAGATCTGAGAACTTAAATGAAGAGA 900
DB 841 AAATCTTGTGCAACATGAGTGGCTCCATGAGAAAGATCTGAGAACTTAAATGAAGAGA 900
QY 901 TATCAAAATCTCTGAGAGACTGGCTTACCAAGAGCTGACGCTTTCAAGAGCCTTAAT 960
DB 901 TATCAAAATCTCTGAGAGACTGGCTTACCAAGAGCTGACGCTTTCAAGAGCCTTAAT 960
QY 961 GTCACTTTGGATTTCTGTGATGTAATAATGAGTCACTGACCTGATGAGAGGGCTTCATAGT 1020
DB 961 GTCACTTTGGATTTCTGTGATGTAATAATGAGTCACTGACCTGATGAGAGGGCTTCATAGT 1020
QY 1021 TTCTGTGTAAGAAAGATTAATGCTTCAATGAAGATCGGGTGACAAAGGCTATGAGAAATG 1080
DB 1021 TTCTGTGTAAGAAAGATTAATGCTTCAATGAAGATCGGGTGACAAAGGCTATGAGAAATG 1080
QY 1081 AAATCTGCCAAGATTAATGCTGCAAGAGAGCTGAGAGCTTTTCAAGCAGCACTGCG 1140
DB 1081 AAATCTGCCAAGATTAATGCTGCAAGAGAGCTGAGAGCTTTTCAAGCAGCACTGCG 1140
QY 1141 ACCAATCAGACCGCTTAAACGGAAGAGACTTGGATTAACAGCAAGAGCAGCTGCG 1200
DB 1141 ACCAATCAGACCGCTTAAACGGAAGAGACTTGGATTAACAGCAAGAGCAGCTGCG 1200
QY 1201 AACCAAGAAACAAAGGCTGTTGGAAGAAAGAAATTAATCTTGGATGATGATACACTA 1260
DB 1201 AACCAAGAAACAAAGGCTGTTGGAAGAAAGAAATTAATCTTGGATGATGATACACTA 1260
QY 1261 CGACTACGAAAGCAGCGGTGATGATCACTTGCCTGATATTTAACTCCCTGTTTTTA 1320
DB 1261 CGACTACGAAAGCAGCGGTGATGATCACTTGCCTGATATTTAACTCCCTGTTTTTA 1320
QY 1321 ACTCAGACCTTTGGTGAAGTTGGCCATGTTCAAGCTGGGGTAGTTAGTTGTTTGG 1380
DB 1321 ACTCAGACCTTTGGTGAAGTTGGCCATGTTCAAGCTGGGGTAGTTAGTTGTTTGG 1380
QY 1381 AAGAGATGTTGTTACCAAGTAAACAACTTATCGCTGTTTACTTGTGCTTTGAA 1440
DB 1381 AAGAGATGTTGTTACCAAGTAAACAACTTATCGCTGTTTACTTGTGCTTTGAA 1440
QY 1441 GTATGATGCGCAGTAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAA 1478
DB 1441 GTATGATGCGCAGTAAATGAAAAAATGAAAAAATGAAAAAATGAAAAAATGAAAAA 1478

RESULT 2

AAA27923 ID AAA27923 standard; cDNA; 1463 BP.

XX AAA27923:

XX 12-SEP-2000 (first entry)

XX Maize Rad2/FEN-1 cDNA.

XX Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.

[illegible]

QY	3119	TTTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTTTGATGCGCAAGCCTCTCGATATGAAG	378
Db	307	TTTACTGGAAGCGGGGAATCAAGCCAGTTTATGTTTTTGATGCGCAAGCCTCTCGATATGAAG	366
QY	3719	AAACGAAGAACTTGGTAAAGATATCTCAAAAAGAGATGATCAACCAAAATGTCATGAG	438
Db	367	AAACGAAGAACTTGGTAAAGATATCTCAAAAAGAGATGATCAACCAAAATGTCATGAG	426
QY	4319	GCAGTAGAGGTAGAGATTAAGATGCGATTGAAAAAATTGAGCAAGAGACTGTAAAGGTC	498
Db	427	GCAGTAGAGGTAGAGATTAAGATGCGATTGAAAAAATTGAGCAAGAGACTGTAAAGGTC	486
QY	4919	ACAAGGCAACCAACGAAGATTTGTAACGACTTTTAAGACTTATGGGGGTCCTGTTGTA	558
Db	487	ACAAGGCAACCAACGAAGATTTGTAACGACTTTTAAGACTTATGGGGGTCCTGTTGTA	546
QY	5519	GAGGCACCTTCTGAAGCAAGAACGAAATGTGCAGCCCTTTGCTTAACATTAAGGTGTC	618
Db	547	GAGGCACCTTCTGAAGCAAGAACGAAATGTGCAGCCCTTTGCTTAACATTAAGGTGTC	606
QY	6119	GCTGTGCTTTAGAAGATTAAGGACTCCCTTACTTTTGGGGCTCCACGGTTCTTGTCAT	678
Db	607	GCTGTGCTTTAGAAGATTAAGGACTCCCTTACTTTTGGGGCTCCACGGTTCTTGTCAT	666
QY	6719	TTAATGATCCCAATGCTCAAGAAATACCTGTATGGAATTTGATTTGCTCCCAAGGTTTG	738
Db	667	TTAATGATCCCAATGCTCAAGAAATACCTGTATGGAATTTGATTTGCTCCCAAGGTTTG	726
QY	7319	GAGGAGCTTGAACCTCACCATGAGACCAATTCATTGATTTTGCACTCCTGTTGATGTAC	798
Db	727	GAGGAGCTTGAACCTCACCATGAGACCAATTCATTGATTTTGCACTCCTGTTGATGTAC	786
QY	7919	TATTGTGATFACCATCAAGGATATCGGGGGGCAACAGCTGTGAACCTTTGTCACAT	858
Db	787	TATTGTGATFACCATCAAGGATATCGGGGGGCAACAGCTGTGAACCTTTGTCACAT	846
QY	8519	GGGTCCATAGAAGCATCTTGAGAAATCTTTAATTAAGACAGATATCAATTTCTCAGAGC	918
Db	847	GGGTCCATAGAAGCATCTTGAGAAATCTTTAATTAAGACAGATATCAATTTCTCAGAGC	906
QY	9119	TGGGCTTACCAAGAAAGCTGCAGGCTTGTTCAAGAGCCCTAATGTCACTTGGATATTCCT	978
Db	907	TGGGCTTACCAAGAAAGCTGCAGGCTTGTTCAAGAGCCCTAATGTCACTTGGATATTCCT	966
QY	9719	GAGCTAAATATGACTGCAGCTGTATGAGAGAGGCTCTATTAAGTTTCTCGTAAAGATAT	1038
Db	967	GAGCTAAATATGACTGCAGCTGTATGAGAGAGGCTCTATTAAGTTTCTCGTAAAGATAT	1026
QY	10319	GCTTTTCATGAAGATCGGGGTGACAAAGGCCATTAAGAAATCAATCTGCGAAGATTA	1098
Db	1027	GCTTTTCATGAAGATCGGGGTGACAAAGGCCATTAAGAAATCAATCTGCGAAGATTA	1086
QY	10919	TCGTGCGAAGGAACATCGAGTCTCTTTTTCACCAACATGCGCACCAATCACACCGGTA	1158
Db	1087	TCGTGCGAAGGAACATCGAGTCTCTTTTTCACCAACATGCGCACCAATCACACCGGTA	1146
QY	11519	AAACGGAAGAGACTTCGATTAAGCAAGCAAGCGCTGCGACAGCAAGAAAACAAAGCT	1218
Db	1147	AAACGGAAGAGACTTCGATTAAGCAAGCAAGCGCTGCGACAGCAAGAAAACAAAGCT	1206
QY	12119	GCTGGAAGGAAGAAATATCTGATGCTTGATGTCAACTGACACTGACACTGCAAGACGGG	1278
Db	1207	GCTGGAAGGAAGAAATATCTGATGCTTGATGTCAACTGACACTGACACTGCAAGACGGG	1266
QY	12719	TGGCATGATCACTTCGCTGATTTATTTAACTCCCTGTTTAACTCAGACCTCTGGTGAA	1338
Db	1267	TGGCATGATCACTTCGCTGATTTATTTAACTCCCTGTTTAACTCAGACCTCTGGTGAA	1326
QY	13319	AGTTGGCCATGTTTCAAGCTGGGGGTAAAGTATGTTGTGTTTAAAGATTTGTTGACCA	1398
Db	1327	AGTTGGCCATGTTTCAAGCTGGGGGTAAAGTATGTTGTGTTTAAAGATTTGTTGACCA	1386
QY	13919	GTAACAAACCTATACGCTGTTTTTAACTCTGCTGCTTGAAGATATGATCGCATTA	1458

Db 1387 GTACAAACTTTCGCTTTTCTTCTCTCTCTTGAAGTAAAAA 1446
Oy 1459 AAAAAAAAAAAAAA 1475
Db 1447 AAAAAAAAAAAAAA 1463
RESULT 3
AAAZ7924
ID AAAZ7924 standard; cDNA; 1541 BP.
XX
AC AAAZ7924;
XX
DT 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 cDNA.
XX
KM Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX
OS Zea mays.
FH Key Location/Qualifiers
FT CDS 79..1218
FT /tag= a
XX
PN W0200036109-A1.
XX
PD 22-JUN-2000.
XX
PF 16-NOV-1999; 99WO-US27147.
XX
PR 15-DEC-1998; 98US-0112332.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Mahajan PB;
XX
DR WPI: 2000-452026/39.
DR P-PADB: AAY95308.
XX
XX
PT Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX
PS Example 1; Page 73-74; 85bp; English.
XX
XX The present sequence is that of maize cDNA (ATCC PTA-533) coding
CC for RAD2/FEN-1 (see AAY95308). The corresponding RNA was isolated
CC from a B73 line seedling after a 10 day drought, heat shocked for
CC 10 hr, and allowed to recover under normal conditions. Rad2/FEN-1
CC is a structure specific endonuclease which under certain conditions
CC also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used
CC to produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX Sequence 1541 BP; 473 A; 308 C; 377 G; 383 T; 0 other;
Query Match 94.8%; Score 1401.2; DB 21; Length 1541;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1409; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Oy 33 GCGGTTCTTGGCCACTCCGGCTCAGCCGCGCCGCCACCCGCCAGCCGCCGAG 92
Db 15 GCGGTTCTTGGCCACTCCGGCTCAGCCGCGCCGCCACCCGCCAGCCGCCGAG 74
Oy 93 CGAGATGGGCATCAAGGGTTTGACGAACCTGCTGGCGACATGCGCCAGCGATGAA 152
Db 75 CGAGATGGGCATCAAGGGTTTGACGAACCTGCTGGCGACATGCGCCAGCGATGAA 134
Oy 153 GGAGCAGAGATTCGAGAGCTACTTGGCCGGAATATGCGCTGAGCGCCAGATGAGCAT 212
Db 135 GGAGCAGAGATTCGAGAGCTACTTGGCCGGAATATGCGCTGAGCGCCAGATGAGCAT 194
Oy 213 CTACAGCTTCTGATAGTACTTGGAGAGACAGCATGGAACCTCTCAATGAGCTGG 272
Db 195 ATACCACTTCTGATAGTACTTGGAGAGACAGCATGGAACCTCTCAATGAGCTGG 254
Oy 273 TGAAGTCATAGTCAATTCGACGAAGATGTTCAACCGACATTAAGATTAAGTACGAGCGGG 332
Db 255 TGAAGTCATAGTCAATTCGACGAAGATGTTCAACCGACATTAAGATTAAGTACGAGCGGG 314
Oy 333 AATCAGCCACTTTATGTTTTGATGGCAAGCCCTGATATGTAAGAAACAGACTTGC 392
Db 315 AATCAGCCACTTTATGTTTTGATGGCAAGCCCTGATATGTAAGAAACAGACTTGC 374
Oy 393 TAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTGCTGAGGACAGTATAGG 452
Db 375 TAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTGCTGAGGACAGTATAGG 434
Oy 453 AGATTAAGATCCGATTTGAAAAATTTGACGACAGAGATCTGTAAGTGTCAACAGCACCAA 512
Db 435 AGATTAAGATCCGATTTGAAAAATTTGACGACAGAGATCTGTAAGTGTCAACAGCACCAA 494
Oy 513 CGAATGTTTAAAGACATTAAGATTAAGGAGGTTCTGTTAAGGACCTTGTGA 572
Db 495 CGAATGTTTAAAGACATTAAGATTAAGGAGGTTCTGTTAAGGACCTTGTGA 554
Oy 573 AGCAGAACAGATATGTCAGCCCTTTGATTAAGGATTAAGTGTTCCTGTTCTTGA 632
Db 555 AGCAGAACAGATATGTCAGCCCTTTGATTAAGGATTAAGTGTTCCTGTTCTTGA 614
Oy 633 AGATTAAGATCCCTTACTTTTGGGGCTCCAGCGTCTCTGCTATTAATGATTCAG 692
Db 615 AGATTAAGATCCCTTACTTTTGGGGCTCCAGCGTCTCTGCTATTAATGATTCAG 674
Oy 693 TTCCAGAAAAATACCTGTGATGAGAAATTTGATGTTGCCAAGGTTTGGAGAGCTTGA 752
Db 675 TTCCAGAAAAATACCTGTGATGAGAAATTTGATGTTGCCAAGGTTTGGAGAGCTTGA 734
Oy 753 CACCATGAGACAGTTCAATGATTTGTGATCCTGTGTGATGATGATGATGATGAT 812
Db 735 CACCATGAGACAGTTCAATGATTTGTGATCCTGTGTGATGATGATGATGATGAT 794
Oy 813 CAAAGGTATCGGGGGGCAACAGCTCTGAACCTTATTCGTAACATGGGTCATAGAA 872
Db 795 CAAAGGTATCGGGGGGCAACAGCTCTGAACCTTATTCGTAACATGGGTCATAGAA 854
Oy 873 CATCTTGGAGATTTTAATAAGACAGATATCAATTCCTAGAGACTGGCTTACAGA 932
Db 855 CATCTTGGAGATTTTAATAAGACAGATATCAATTCCTAGAGACTGGCTTACAGA 914
Oy 933 AGCTGAGCGTTGTTCAAGAGACCTTAATGTCATTTGATGATTTGATGATGATGAT 992
Db 915 AGCTGAGCGTTGTTCAAGAGACCTTAATGTCATTTGATGATTTGATGATGATGAT 974
Oy 993 TGCACCTGATGAGGAGGCTCATAGTTTCTGTTAATAATATGTTTCAATGAGA 1052
Db 975 TGCACCTGATGAGGAGGCTCATAGTTTCTGTTAATAATATGTTTCAATGAGA 1034
Oy 1053 TCGGGTGACAAAGGCCATAGAGAAATCTCCAGAGATTAATGTTGCAAGAGAG 1112
Db 1035 TCGGGTGACAAAGGCCATAGAGAAATCTCCAGAGATTAATGTTGCAAGAGAG 1094


```
QY 901 TATCAATTCCTGAGACGTGGCCTTACCAAGAGCTGAGCGTTGTTCAAGAGCCTAAT 960
DB 841 TATCAATTCCTGAGACGTGGCCTTACCAAGAGCTGAGCGTTGTTCAAGAGCCTAAT 900
QY 961 GTCACATTCGATATTCCTGAGCTGAGAAATGAGCTGACCTGATGAGAGGCTCATTAAGT 1020
DB 901 GTCACATTCGATATTCCTGAGCTGAGAAATGAGCTGACCTGATGAGAGGCTCATTAAGT 960
QY 1021 TTCCTGTAAAGATATATGTTTCATGAAAGATCGGGTGTGCAAAAGCCCTAGAGAAATG 1080
DB 961 TTCCTGTAAAGATATATGTTTCACGAGATCGGGTGTGCAAAAGCCCTAGAGAAATG 1020
QY 1081 AATCTCTCAAGAAATATATGTCGCAAGAGAGCTGAGTCTTTTCAAGCCAACTGCC 1140
DB 1021 AATCTCTCAAGAAATATATGTCGCAAGAGAGCTGAGTCTTTTCAAGCCAACTGCC 1080
QY 1141 ACCACATCAGCAGCGCTAAACGAGAGAGCTTCGATTAACAAAGAGAGCAGCTGGC 1200
DB 1081 ACCACATCAGCAGCGCTAAACGAGAGAGCTTCGATTAACAAAGAGAGCAGCTGGC 1140
QY 1201 AACAGAAACAAAGAGCTGGGAGAGAGAGAAATATCTTGATGCTGTGATACAACTA 1260
DB 1141 AACAGAAACAAAGAGCTGGTGTGAGAGAGAGAAATATCTTGATGCTGTGATACAACTA 1200
QY 1261 CGACTACGAAAGAGCAGCGTGGCATGATCACTGCGCTAGATTAATTAACCTCCCTGTTTA 1320
DB 1201 CGACTACGAAAGAGCAGCGTGGCATGATCACTGCGCTAGATTAATTAACCTCCCTGTTTA 1260
QY 1321 ACTAGACCTTTGGTGAAGATTTGCCATGTTTCAAGCTGGGGTAAATAGTTGTGTG 1380
DB 1261 ACTAGACCTTTGGTGAAGATTTGCCATGTTTCAAGCTGGGGTAAATAGTTGTGTG 1320
QY 1381 AAGAGATTTGGTGTACCAAGTAAACAACTTATCGCTGTTT 1423
DB 1321 AAGAGATTTGGTGTACCAAGTAAACAACTTATCGCTGTTT 1383

RESULT 5
AA02117
ID AA02117 standard; DNA; 1144 BP.
XX
AC AA02117;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 DNA fragment.
XX
KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; ss.
XX
OS Homo sapiens.
XX
PN US874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR.
XX
DR WPI: 1999-179985/15.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
PS e.g. recombinant polypeptides
XX
PS Claim 3; Column 25; 58pp; English.
XX
```

```
CC This sequence encodes a human FEN-1 (flap endonuclease) protein. This
CC protein can be used in methods for detecting a pathological condition in
CC a patient, for diagnostic purposes, for screening for antineoplastic
CC agents and carcinogens, for diagnostic staging of neoplasia, for
CC producing recombinant flap endonuclease for use as research or
CC diagnostic reagents, for producing antibodies reactive with the novel
CC polypeptides, for producing transgenic nonhuman animals expressing the
CC novel polypeptides encoded by a transgene. The invention also provides
CC novel molecular cloning techniques and reagents involving cleavage of
CC a flap or nick with a flap endonuclease.
XX
SQ Sequence 1144 BP; 292 A; 283 C; 353 G; 216 T; 0 other.
```

Query Match 20.3%; Score 299.4; DB 20; Length 1144;

Best Local Similarity 56.5%; Pred. No. 4.3e-70;

Matches 599; Conservative 0; Mismatches 456; Indels 6; Gaps 2;

```
QY 97 ATGSCATCAAGGGTTTACGAACTGCTGGCGGCAATGCGCCCAAGCGCATGACAGAG 156
DB 1 ATGGAATTCAGAGCGCTGCGCAAACTAATTCGTATGTGGCCCGCCAGTCCGAGGAG 60
QY 157 CAGAGTTCGAGAGCTACTGCGCGCAAAATCGCGTCGAGCGCAGCATGAGCATCTAC 216
DB 61 AATGACATCAAGAGCTACTTTGGCCGTAAAGTGGCCATGATGCTCTATAGCATTTAT 120
QY 217 CAGTTCCTGATAGTATGTTGAGAGAGAGCATGGAACCTCTACAAATGAAGCTGTGAA 276
DB 121 CAGTTCCTGATAGTCTGTTCGCGAG--GCTGGGATGTGCTGCAAGAAATGAGAGAGT 177
QY 277 GTCACTAGCATTTTCAAGGAATGTTCAACCGGACAAATAGATTACTGTGAAGCGGAAATC 336
DB 178 ACCACAGCCACCATGATGGGCAATGTTTACCCACATTCGATGATGAGAAACGGCATC 237
QY 337 AAGCCAGTTTATGTTTGTGATGCAAGCCTCCTGATTAAGAAACAAAGACTGTTAA 396
DB 238 AAGCCGATGATGCTTTGATGAGCAAGCCGCGACAGCTCAAGTCAAGGAGACTGGCCAA 297
QY 397 AGATACCTCAAAAAGAGATGAGCAACCAAGATCTGACTGAGAGAGTAAAGAT 456
DB 298 CCGAGTGAAGCGGCGCTGAGAGCAGAGACGTCGACAGGCTCAGGCTGTGGGGCC 357
QY 457 AAGATGCTATGAAAAATGAGCAAGAGAGCTGTAAGTCAACAGGCAACACAGCA 516
DB 358 GAGCAGAGAGTGGAAAAATCTAAGCGGCTGTGAAGGCTCACTAAGCAGCAGCATAT 417
QY 517 GATTGAACGACATTAAGACTTATGGGGGTTCCTGTTGAGAGAGCACTTCTGAAGCA 576
DB 418 GAGTGAACATCTGCTGAGCCTCATGGGCATCCCTTATCTTGATGACCCAGTGAAGCA 477
QY 577 GAGCAGAAATGTGACGCCCTTTGCAATAACGATTAAGTGTGCTGTTCAGAGAT 636
DB 478 GAGGCAGCTGTGCTGCTGCTGCTGCAAGGCTGCAAGCTATGCTGCGCTACCGAGAC 537
QY 637 AAGGACTCCCTTACTTTGGGGCTCCAGGCTCCCTGCTGATTAATGATGATCAAGTTCC 696
DB 538 ATGAGCTGCTCACTTGGCAGCCCTGTGCTAATGCAACATCTGACCTGAGAGCC 597
QY 697 AAGAAATACCTGTGATGAATTTGATGTTGCCAAGTTTGGAGAGCTGGAATCACTCAC 756
DB 598 AAAAGAGCTCAATCCAGGAATTCACACCTGAGCGGATTTGACAGAGCTGGGCTGAA 657
QY 757 ATGAGCAATTCATGATTTGTCATCTGTGTGATGATGACTATTTGATAGATCAAA 816
DB 658 CAGGAACAGTTTGTGATCTGTGATCTGCTGATGAGCAGTACTGTGAGATATCCGG 717
QY 817 GGTATCGGGGGGCAAGACCTGTAACCTTATTCGAACATGGGTGATGAAGAAAGCATC 876
DB 718 GGTATTTGGGCCCAAGCGGCTGTGACCTCATCCAGGAAGCAAAAGATGAGAGATC 777
QY 877 TTGAGAAATCTTAATTAACAGATATCAATTCCTGAGAGACTGGCTTACCAAGAACT 936
DB 778 GTGCGGCACTTGACCCCAACAGTAACTGTGCGAAGAAATTTGGCTCCACAGAGAGCT 837
```


PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 XX Disclosure; Fig 3B; 58bp; English.
 PS
 XX

CC This sequence encodes a yeast FEN-1 (flap endonuclease) protein. This
 CC protein is used in a method to isolate novel human FEN-1 proteins for
 CC detecting a pathological condition in a patient, for diagnostic purposes,
 CC for screening for antineoplastic agents and carcinogens, for diagnostic
 CC staging of neoplasms, for producing recombinant flap endonuclease for use
 CC as research or diagnostic reagents, for producing antibodies reactive
 CC with the novel polypeptides, for producing transgenic nonhuman animals
 CC expressing the novel polypeptides encoded by a transgene. The invention
 CC also provides novel molecular cloning techniques and reagents involving
 CC cleavage of a flap or nick with a flap endonuclease.
 CC
 XX
 XX Sequence 1149 BP; 410 A; 182 C; 279 G; 278 T; 0 other:

Query Match 18.0%; Score 265.4; DB 20; Length 1149;
 Best Local Similarity 55.28; Pred. No. 5.7e-61;
 Matches 596; Conservative 0; Mismatches 456; Indels 27; Gaps 3;

QY 97 ATGGGATCAAGAGGTTTGACGAACTGCTGGCGGACATGCGCCGACAGGCGATGAGGAG 156
 DB 1 ATGGGTATTAAGGTTTGATGACATTTATGCGAACATGTTCCCTGCTATGACGAA 60
 QY 157 CAGAGTTGAGAGCTACTTGGCGCAAAATCGCCGTGACGCCGACGATGACATCTAC 216
 DB 61 AGCGATATCAAGAGCTTTTGGCAGAAAGTTGCCATGCAATGCTCATATAT 120
 QY 217 CAGTTCCTGATAGTGTGSAAGGACAGCATGAACTGCACAAATGAACTGTGAA 276
 DB 121 CAGTTTAAATGCTGTAAGACAGCAAGACGTTGGCAGTTGACCAATGAACCCGCTAA 180
 QY 277 GTCACTACTGTTTGCAGAAATGTTCAACCGACATTAAGATTCTGGAAGCGGAATC 336
 DB 181 ACAAGTCACACTGATGGGTATGTTTATAGACACTGAGATATGATTAACGCTATC 240
 QY 337 AAGCAGATTATGTTTTGATGCAAGCCTCTGATATGAAGAAACAAGATTCCTAA 396
 DB 241 AAGCCTGTATGTTCTGACGCAACCTCAGCTTGAATCTCATGAGTTGACAAAG 300
 QY 397 AGATACTCAAAAAGATGATGCAACCAAGATCTGACTGAGGCGATGAGTAGAAT 456
 DB 301 CGGCTTCAAGAGGCTGGAACGAAAAAACTGCGAGAGCAACAAGAAATTGAA 360
 QY 457 AAAGATGCGATTGAAAAATTGACAGAGGACTGTAAGTCACAAAGCACACAAACGA 516
 DB 361 AAGATGAAGCAAGAAAG-----AGATTGTTGAAGTCTCAAGAGCATTAATGA 411
 QY 517 GATGTAAAGCACTATTAGACTATGAGGCTTCCTGTTGAGAGGACCTTCTGAACGA 576
 DB 412 GAAGCCCAAAATTAAGTAAGACTAATGGAATCCCATATATAATGAGCCCAAGCAACT 471
 QY 577 GAACAGAAATGTCAGGCCCTTGATTAAGATAAGTGTGCTGCTGCTTCAAGAAAT 636
 DB 472 GAGGCTCAATGCTGAGATTGGCAAGAGGAAAGGTGATGCGCGCAAGTGAAGAT 531
 QY 637 AAGGACTCCCTTACTTTGGGGCTCCACGGTTCCTGCTCATTTAATGATCAAGTTCC 696
 DB 532 ATGACACACACTCTGTATAGAACACCTCTTGTGAGACATTTGACATTTTTCAGAGCC 591
 QY 697 AAGAAATACCTGTGATGAGATTTGATGTTCCCAAGTTTGGAGAGGACTTGAACCTCAC 756
 DB 592 AAGAAAGAACCGATTCAAGAAATAGATAGTAATAGTTAGAGAGGACTGACCTTGACA 651
 QY 757 ATGACCAAGTTGATGATTTGTGATCCTGATGATGATGATGATGATGATGATGATG 816
 DB 652 ATAAAGAGAGTTGTTGATCTTGTGATATGCTGTGTTGTTCTCTGTTGAAGATGACA 711
 QY 817 GGTATCGGGGGGCAAGACTCTGAACCTTATTCGTCACAAATGAGGATGATGAAAGCATC 876
 DB 712 GGTGTTGGTCCAGTACAGCCTTAAATTAATGAATAAACGATGATCCATGCAAAATATC 771

QY 877 TTGAGATCTT-----AATTAAGACATATCAATTCGTGAGACCTGG 921
 DB 772 GTGGACTTATTTGATCTTGGGAGTCGAACACACTAAATGGAATCCGACAGACTGG 831
 QY 922 CCTTACCAAGAACTGACGCTTGTTCAGAGAGCCCTAATGCTACTTTGGATA---TTCTT 978
 DB 832 CCTTACCAAGAACTGACGCTTGTTCAGAGAGCCCTAATGCTACTTTGGATA---TTCTT 978
 QY 979 GAGCTTAAATGACGCTGACCTGATGAGAGGCTTCTCATAGTTTCTGCTGTAAGATTAAT 1038
 DB 892 AACTTGAATGCTGCGCCACCAAGGAGAACTTATGAGTATTTATGTATGATGATAG 951
 QY 1039 GGTTCATGAGATCGGGTGACCAAGCCATAGAGAAATCAATTCGCAAGATTA 1098
 DB 952 AATTCAGTGAAGAAAGATTTAATCTGTTATATCAAGATTGAATAAAGCTTGAATCT 1011
 QY 1099 TCGTCGCAAGAAAGCTGACGCTTGTTCAGAGAGCCCAACTGCCACATCAGCAGCT 1157
 DB 1012 GGCATTCAAGGTAAGTTAGATGAGTTCTTCCAAAGTGCTCTTAAGCAAGAGACGCT 1070

RESULT 11
 ABL14286
 ID ABL14286 standard; cDNA; 3362 BP.
 XX
 XX ABL14286;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID. NO 37340.
 KW Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 PD 27-SEP-2001.
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PL Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR P-PSDB: ABB70183.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS
 PS Claim 1; SEQ ID NO 37340; 21pp + Sequence Listing; English.
 PS
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pt_sequences.
 CC
 XX
 XX Sequence 3362 BP; 979 A; 709 C; 750 G; 924 T; 0 other;

Query Match	17.0%;	Score 251;	DB 23;	Length 3362;
Best Local Similarity	55.1%;	Pred. NO. 7e-57;		
Matches 535;	Conservative 0;	Mismatches 430;	Indels 6;	Gaps 2

OY	178	GGAGCGAAATCCCGTCAGCGCACATGACATCTCAATTCCTATAGTAA	237
Db	1144	GGTCGCAAGGTACCAATCCATGCTAGTATGTGCTGTACCAATTCCTATAGTAA	1203
OY	238	AGGACAGGCACTGAAACCTCCACAAATGAAGCTGGTGAATCATTGATTTGCAAGA	297
Db	1204	TCCGCAAGGCG---GCCACGTTGGCGCACCGTAAATGGTGTGATCCCACTCCCACTTAATGGGC	1260
OY	298	ATGTTCAACCGGCAATTAAGATTAATCTGGAAGCGGGGAATCAACCGCTTATGTTTGCAT	357
Db	1261	ATGTTTATCCCGCACTCCGATTTGCTGTGAGCAACAGATCAACCGGGTATATGTCTTGAC	1320
OY	358	GGCAGGCTCTGTATATGAAGAAACAAGACTTGCTAAAGATACTCAAAAAGATGAT	417
Db	1321	GGAAAGCCACCAATCTCAAGTCCGGGTGACTGGGCAAGCGCGGCAGCGCGGAGAGA	1380
OY	418	GCAACCAAAATGTCAGTGGAGCAATGAAGGTAAAGATGCGATTGAAAAATVG	477
Db	1381	GGGAGAGAGGCACTGAGAGCGCGCCACCAGTGGCGGAGATGATCCGGTATGAAAAGCTTT	1440
OY	478	AGCAAGAGGACTGTAAAGGTACACAAGGCAACAACGAAGATTGTAAAGCACTTAAGA	537
Db	1441	AATGCGGATTTGGTCGGGGTACGAAGAGAGACGCCCAAGAGAGCCAAAGAACTGCTACA	1500
OY	538	CTTATGGGGGTTCTGTGTGAAGGCAACCTTCGAAAGCAAGAGATGAGAGCCCTT	597
Db	1501	CTAATGGGTGGCCCTATGTGTGATGCAACCGTGGAGAGCGGAGCCCATGTGCGCTTG	1560
OY	598	TGCATTAAGCATAAAGTGTGCTGCTGTGCTTGAAGATTAAGGACTCCCTTACTTTGGG	657
Db	1561	GTGAAGACTGGAAAGGTTTATGCCACCGCACGAGAGATATGATGCCCTACATTTGCGA	1620
OY	658	GCTCCAGGTTCCCTGCTCATTTAAATGATTCGAAGTCCCAAGAAATACCTGTGATGAA	717
Db	1621	TCTCAAAACTGTTGAGATACCTTACCCTACAGGAGGACGAAAGATGCCCTCAAGAG	1680
OY	718	TTTGATGTGCCAAGGTTTGGAGAGGCTTGAACCTCACATGAGCACTTCAATTTGAT	777
Db	1681	TTACGCTACGACAAGCTGTGGGAAGGTCTGGCCATTACAAATCGAGAGTTCAATGATCTA	1740
OY	778	TGCATCTGTGTGGATGTGACTATGTGAATGATCAAGAGTATGGGGGCAACCACT	837
Db	1741	TGTTATTCTGCTGGGTTGCCGTTACTGTGAGAGCATTAAGGATTTGGACCCCAAGGAGCG	1800
OY	838	CTGAACCTTATTCCTCAACATGGGTCATAGAAGACATCTGGAGATCTTAATAAAGAC	897
Db	1801	ATCAGACTGATCAACACCTATCGGGATATGAGACATATCTGGATTAACCTGGACCTAGC	1860
OY	898	AGATATCAAAATTCCTGAGGAGTGGCCCTACCAAGAGCTGCAGACCTGTTCTCAAGGAGCT	957
Db	1861	AAATPACACCGTGGCCGGAATCTGAATCAAGAGTGGCCGCGGAATCTTCAATCCAACCG	1920
OY	958	AATGTACAC---TTTGATATTCCTGTGAGCTAAATGTGACATGCACTGTATGAGAGGAGCTTC	1014
Db	1921	GAGGATGCTGATGCGCATCTCATATGATCTCAATATGGGTGAGCGCGGATGAGAGGGCTT	1980
OY	1015	ATAAGTTTCCCTGGTAAAAAGATTAATGTTTCAATGAAGATCGGGTGCACAAAGCCATAGAG	1074
Db	1981	GTCAAGTTTCTCTGGCGGACCGGCACTTCACGAAGACGCGGCTTCGAACGCTGCCAAA	2040
OY	1075	AAGATCAAAATTCGCGAAGATTAATGCTGCAAGAGAGCACTCGAGTCCCTTTCAAGCCA	1134
Db	2041	AAGGTATGAATATCCAGAGCGGCCAGAGCTCAGGTAGACTGATATAGCTTTTAAAGCA	2100
OY	1135	ACTGCCACAC 1145	
Db	2101	CTGCCACAGCA 2111	

RESULT 12

ABL10094
ID ABL10094 standard; cDNA; 5471 BP.
xy

AC ABL10094;

DT 26-MAR-2002 (first entry)

DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 24764.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

05 *Drosophila melanogaster*.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 200005-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75

DR P-PSDB; ABB65991.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

PS Claim 1; SEQ ID NO 24764; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention
CC is useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB161176-AB130511), expressed DNA
CC sequences (AB101840-AB161175) and the encoded proteins
CC (AB573737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WFO
CC at http://wipo.int/pub/published_pat_sequences.

Sequence 5471 BP; 1422 A; 1223 C; 1325 G; 1501 T; 0 other;

Query Match	11.7%	Score 173.2	DB 23	Length 5471
Best Local Similarity	54.0%	Pred. NO. 6.7e-36		
Matches 377	Conservative	0	Mismatches 318	Indels 3
				Gaps 1

451 GGAGTAAGATGCCATTTGAAAAATTGAGCAAGAGGACTGTAAGGTCACAAGGCAACAC 510

Db 4 GGAGATGATGCCCGTATCGAAAGTTAATCGCCGATGGTCCGGGTAACGAAGGAGCAC 63

511 AACGAGATTGTAACGACTATTAGACTTATGGGGTTCCTGTTGTAGAGGCACCTTCT 570

Db 64 GCCAAGAGGCCAAGGACTGCTCACACTAATGGGTGCGCCCTATGTTGATGCACCGTGC 123

571 GAAGCAGAAGCAGATGTGCAGCCCTTTGCATTAACGATAAGGTTCGCTGTTGCTTCA 630

Db 124 GAAGCGAGGCCACGTGTGCCGCTCTGCTGAAGCTGGAAGGTTATGCCACCGCCACG 183

631 GAAGATAAGGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGTCATTTAATGATCCA 690

Db 184 GAGGATATGGATGCCCTCACATTCGGATCTACAAACTGTTGAGATACCTTACCTACAGC 243

691 AGTCCAGAAATACCTCTGATGGAATTGATGTTGCCAAGGTTTGGAGGAGCTTGAA 750

```
Db 244 GAGGACCAAGATGCCCGTCAAGAGATTACGCTACGACAAAGCTGTGGAGCTGTGGCC 303
Qy 751 CTCACCATGGAACAGATTGATTTGTCATCTGTGTGATGTGATTTGATAGC 810
Db 304 ATTACATTCGAGAGTTCATTTATGATTTGCTGCGGTTGGGATTTACTGAGAC 363
Qy 811 ATCAAGGATTCGGGGGCAACAGCTCTGAAACTTATTGCTCAACATGGCTCATAGA 870
Db 364 ATCAAGGGTATTGGACCCAGCGATCGAACCTGATCAACACCTATCGGATATAGG 423
Qy 871 AGCATCTGGAGAAATCTTAATAAGACAGATATCAATTTCCAGAGACTGGCTTACCA 930
Db 424 ACTATTCTGGATTAACCTGACTGACTAGCAATACACCTGCGGAGAACTGGAACTAAG 483
Qy 931 GAACCTCGACGCTTGTTCAGAGAGGCTATGCTCACTTTGATATTCCT---GAGTAA 987
Db 484 GTGGCGCGGGAACCTTTCATGGAACCGAGAGTGTGATGCGGACCTCATATGATTC 543
Qy 988 TGGACTGACCTGATGAGAGAGGCTCTCATTAAGTTCTGTGTAAGATATATGCTTCA 1047
Db 544 TGGGTGAGCGGATGAGAGAGGCTTGTCAAGTTCTGTGCGGAGCCGCGAGTTCAAC 603
Qy 1048 GAAGATCGGGTGACAAAGCCATAGAGAAGATCAATCTGCAGAGATTAATTCGCGAA 1107
Db 604 GAAGAGCCCGTTCGCAAGCGTGCCAAAAGCTGATGAATCCAAAGCCAGCCAGACTG 663
Qy 1108 GGAAGACTCGAGTCTTTTTCAGGCCAACAATGCCACAC 1145
Db 664 GTGAGACTCGATAGCTTCTTTAAGCACTGCGCAGCAC 701

RESULT 13
AAT76685 standard; DNA: 1023 BP.
XX
AC AAT76685;
XX
DT 14-APR-1998 (first entry)
XX
DE Pyrococcus furiosus FEN-1 endonuclease gene coding sequence.
XX
KM Nucleic acid cleavage: 5' nuclease; DNA cleavage: RNA cleavage;
KW Invader directed cleavage: FEN-1; endonuclease: ds.
XX
OS Pyrococcus furiosus.
XX
PN W09727214-A1.
XX
PD 31-JUL-1997.
XX
PF 22-JAN-1997; 97MO-US01072.
XX
PR 02-DEC-1996; 96US-0759038.
PR 24-JAN-1996; 96US-059491.
PR 12-JUL-1996; 96US-0682853.
PR 29-NOV-1996; 96US-0756386.
PR 02-DEC-1996; 96US-0758314.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichiev VI;
PI Olive DM, Prudent JR.
XX
DR P-PSDB; AAW24216.
XX
PT Thermostable structure-specific nuclease(s) - used for detection and
PT characterisation of nucleic acid sequences and variations in nucleic
PT acid sequences
XX
PS Example 28; Page 283-285; 457pp; English.
XX
CC This sequence comprises the coding region of the gene encoding
```

```
CC Pyrococcus furiosus (Pfu) FEN-1 endonuclease (see AAW24216).
CC It was obtained by PCR amplification (see AAT76682-83). Large
CC scale preparation of recombinant Pfu FEN-1 was performed using
CC E. coli as host. Pfu FEN-1 is a thermostable enzyme. It can be
CC used in novel methods for the detection and characterisation of
CC nucleic acid sequences and variations in nucleic acid sequences.
SQ Sequence 1023 BP; 396 A; 172 C; 240 G; 215 T; 0 other.

Query Match 11.0%; Score 162.8; DB 18; Length 1023;
Best Local Similarity 55.0%; Pred. No. 1.8e-33;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

Qy 146 CATGAGAGAGCAGAGTTCCGAGACTACTTGGCGCGAATAATCCCGCTCGACGCCAGCA 205
Db 29 CAGAAAGAAATTAAGTAGTAAACCTATACGAGAAAAAATTCGCAATTCAGCGTCTTA 88
Qy 206 TGAGCATCTACAGTTCCTCGATAGTAGTTGGAAGAGACAGCATGGAACCTCACAATG 265
Db 89 ATGCAATCTACCAATTTTGTCCACATTAAGACAGAAAGATGGAACCTCATTATGAT 148
Qy 266 AAGCTGTTGAAGTCACTGTGCTATTGCAAGGANTGTTCAACCGGACAAATTAAGATTACTG 325
Db 149 CAAAGGTAAGATTAACCTCCACCTAAGCGGCTCTTTTACAGGACAAATTAACCTAATG 208
Qy 326 AAGCGGAATCAAGCCATTTATGTTTGTATGGCAAGCTCTCTATATGAAGAAACAG 385
Db 209 AGGCTGGAATTAACCTGTGTATGTTTGTATGAGGAACCTCCGAAATTCAGAAAAAGAA 268
Qy 386 AACTTGCTAAAGATCTCAAAAAAGAGATGATGCAACCAAGATGACTGACGAGTAG 445
Db 269 AGCTCGAAAAAAGAGAGAGAGAGAGAGAACTGAAAGAGAGAGAGACACTTG 328
Qy 446 AGTGAAGATTAAGATCGATTGAAAATTGAGCAAGAGACTGTAAAGTCAACAAGC 505
Db 329 AAAAAGAGAGATAGAGAGCAAGAAATATGCCCAAGAGCAACAGGCTAATATGAAA 388
Qy 506 AACACAGAGAAATTTGTAAGCATTTAAGACTATTATGAGGGGTTCTGTTAGAGCGAC 565
Db 389 TGCATCATGAGAGATGCAAAAAAAGCTTTAGACTTTATGGAATTCCTATATGTTCAAGC 448
Qy 566 CTTCTGAAGCAGAGCAGAAATGTGCGAGCCCTTTGCAATTAAGCAATGCTGCTGTTG 625
Db 449 CTACGAGAGGAGAGGCCCAAGCTGCATATATGCGCGCAAGGGAGCGTATATGATCG 508
Qy 626 CTTCAAGAGATTAAGACTCCCTTACTTTTGGGCTCCACGCTTCTTGCTATTAATG 685
Db 509 CTAGTCAAGATTACGATTCCTTACTTTTGGAGCTCCAAAGACTTGTAGAACTTAACAA 568
Qy 686 ATCCAGTTCCAGAAATACCTGTGATGCAATTTGATGTTG 727
Db 569 TAACAGGAAAAAGAAAGTTGCTGGAATAATGTCTACGTG 610

RESULT 14
AAV65840 standard; DNA: 1023 BP.
XX
AC AAV65840;
XX
DT 02-FEB-1999 (first entry)
XX
DE Pyrococcus furiosus FEN-1 endonuclease gene ORF.
XX
KM Nucleic acid detection; multiple sequential invasive cleavage;
KW FEN-1; endonuclease; nuclease: ds.
XX
OS Pyrococcus furiosus.
XX
PN W09842873-A1.
XX
PS 01-OCT-1998.
XX
CC
```


Db 301 TTCAACCGGACATATAGATTAAGTGAAGCGGGAAATCAAGCCAGTTTATGTTTTCATGGC 360
QY 361 AAGCTCTGTATATGAGAAAACAGAACTTGTCTAAAGATATCTCAAAAAAGATGATGCA 420
Db 361 AAGCTCTGTATATGAGAAAACAGAACTTCTTAAAGATCTCAAAAAAGATGATGCA 420
QY 421 ACCAAAGATCTGACGTAGGAGAGTAGAGATTAAGATGCGATGGAATAATGAGC 480
Db 421 ACCAAAGATCTGACGTAGGAGAGTAGAGATTAAGATGCGATGGAATAATGAGC 480
QY 481 AAGAGACTGTAAAGTCAACAGGCACACAGAAATGTTAAAGCATATTAAGACTT 540
Db 481 AAGAGACTGTAAAGTCAACAGGCACACAGAAATGTTAAAGCATATTAAGACTT 540
QY 541 ATGGGGTCTCTGTGTAGAGGACCTCTGAGACAGAGATGTGACCCCTTGGC 600
Db 541 ATGGGGTCTCTGTGTAGAGGACCTCTGAGACAGAGATGTGACCCCTTGGC 600
QY 601 ATAAACGATTAAGGTCGTGCTGCTTCAAGAGATAGGACTCCCTACTTTTGGGCT 660
Db 601 ATAAACGATTAAGGTCGTGCTGCTTCAAGAGATAGGACTCCCTACTTTTGGGCT 660
QY 661 CCAGGTCCTTCTGATTAATGATCAAGTCCAAAGTCCAAAGAAATACCTGTATGGAATTT 720
Db 661 CCAGGTCCTTCTGATTAATGATCAAGTCCAAAGTCCAAAGAAATACCTGTATGGAATTT 720
QY 721 GATGTTGCCAGGTTTGGAGAGGCTTGAATCAACATGAGCCAGTCAATGATTTGGC 780
Db 721 GATGTTGCCAGGTTTGGAGAGGCTTGAATCAACATGAGCCAGTCAATGATTTGGC 780
QY 781 ATCCGTGTGATGATGATATTTGATAGCATCAAAAGTATCGGGGGCAACAGCTCTG 840
Db 781 ATCCGTGTGATGATGATATTTGATAGCATCAAAAGTATCGGGGGCAACAGCTCTG 840
QY 841 AAATCTATCTGCAACATGAGGTCATAGAAAGCATCTTGGAGATCTTAATTAAGACAGA 900
Db 841 AAATCTATCTGCAACATGAGGTCATAGAAAGCATCTTGGAGATCTTAATTAAGACAGA 900
QY 901 TATCAATTCCTGAGAGATGAGCTTACCAAGAGCTCGACGTTTTCAGAGAGCTTAT 960
Db 901 TATCAATTCCTGAGAGATGAGCTTACCAAGAGCTCGACGTTTTCAGAGAGCTTAT 960
QY 961 GTCACTTTGATATCTCTGAGCTAAATGATGATGACCTGATGAGAGGTCCTCAATAGT 1020
Db 961 GTCACTTTGATATCTCTGAGCTAAATGATGATGACCTGATGAGAGGTCCTCAATAGT 1020
QY 1021 TTCTCGTGAAGATATAGTGTTCATATGATCGGTCGACAAAGCCATAGACAGATC 1080
Db 1021 TTCTCGTGAAGATATAGTGTTCATATGATCGGTCGACAAAGCCATAGAGATC 1080
QY 1081 AAATCTGCCAAGATTAATGTCGCAAGAGACTCGAGTCCCTTTTCAAGCCAACTGCC 1140
Db 1081 AAATCTGCCAAGATTAATGTCGCAAGAGACTCGAGTCCCTTTTCAAGCCAACTGCC 1140
QY 1141 ACCACATCAGACCGCTAAAACGGAAGAGACTTCGATTAACAAAGCAAGCAGCTCG 1200
Db 1141 ACCACATCAGACCGCTAAAACGGAAGAGACTTCGATTAACAAAGCAAGCAGCTCG 1200
QY 1201 AAGCAAGAAACAAAGGCTGTGGAAGAAAGAAATATCTTGATCTGTTGATGATGATGAT 1260
Db 1201 AAGCAAGAAACAAAGGCTGTGGAAGAAAGAAATATCTTGATCTGTTGATGATGATGAT 1260
QY 1261 CGACTACGAAAGACGCGTGGCATGATCACTCGCTAGATATTAATTAATTAATTAATTA 1320
Db 1261 CGACTACGAAAGACGCGTGGCATGATCACTCGCTAGATATTAATTAATTAATTAATTA 1320
QY 1321 ACTCAGACCTTGTGTAAGTTTGCATGTTTCAAGCTGGGGTAAGTTAGTTGTTG 1380
Db 1321 ACTCAGACCTTGTGTAAGTTTGCATGTTTCAAGCTGGGGTAAGTTAGTTGTTG 1380
QY 1381 AAGAGATGTGTACCAAGTAAACAAACTTATCGCTGTTTACTTCTGTCTTGA 1440
Db 1381 AAGAGATGTGTACCAAGTAAACAAACTTATCGCTGTTTACTTCTGTCTTGA 1440

QY 1441 GATGATATGCGAGTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1478
Db 1441 GATGATATGCGAGTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1478

RESULT 2
US-09-426-557-1
; Sequence 1, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)...(1221)
US-09-426-557-1

Query Match 96.2%; Score 1421.8; DB 4; Length 1463;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1435; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 19 AATAGCTGCGGTCGCGGTTTCTTCTGCGCACTCCGCGCTCAGCCGCCGCCGCCGCCGCC 78
Db 7 AATAGCTGCGGTCGCGGTTTCTTCTGCGCACTCCGCGCTCAGCCGCCGCCGCCGCCGCC 66

QY 79 ACAGCGCGCGAGAGATGAGGTCATCAAGGTTGAGAAACTCTGCGGAGCAATGCG 138
Db 67 ACAGCGCGCGAGAGATGAGGTCATCAAGGTTGAGAAACTCTGCGGAGCAATGCG 126

QY 139 CCCAAGGCGATGAAGAGCAGAGAGTTCGAGAGCTACTTCGCGCCGCAAAATGCGCGTCGAC 198
Db 127 CCCAAGGCGATGAAGAGCAGAGAGTTCGAGAGCTACTTCGCGCCGCAAAATGCGCGTCGAC 186

QY 199 GCCAGCATGAGCATACCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 258
Db 187 GCCAGCATGAGCATACCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 246

QY 259 ACAATGGAAGCTGGTGAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 318
Db 247 ACAATGGAAGCTGGTGAAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 306

QY 319 TTACTGGAAGGCGGGAATCAAGCCAGTTTATGTTTGTATGCGAAGCCTCTGATATGAG 378
Db 307 TTACTGGAAGGCGGGAATCAAGCCAGTTTATGTTTGTATGCGAAGCCTCTGATATGAG 366

QY 379 AAACGAAGACTTGTAAAGATACCTCAAAAGAGATGATGATGATGATGATGATGATGATGATG 438
Db 367 AAACGAAGACTTGTAAAGATACCTCAAAAGAGATGATGATGATGATGATGATGATGATGATG 426

QY 439 GCAGTAGAGAGTAAAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATG 498
Db 427 GCAGTAGAGAGTAAAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATG 486

QY 499 ACAAGCAACACAAAGATTTAAAGCACTATTAAAGCTTATGAGGTTCTGTTGTA 558
Db 487 ACAAGCAACACAAAGATTTAAAGCACTATTAAAGCTTATGAGGTTCTGTTGTA 546

QY 559 GAGGACCTTCTGAAGCAGAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 618
Db 547 GAGGACCTTCTGAAGCAGAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 606

QY	619	GGGTTGCTCGAAGAAATGAAGACGCCCTTAACCTTTGGGGGCTCCAGCGTTCCTCCAT	678
QY	619	GGGTTGCTCGAAGAAATGAAGACGCCCTTAACCTTTGGGGGCTCCAGCGTTCCTCCAT	678
Db	607	CGTGTGCTTCCAGAAGATATGAGCTCCCTTACTTTTGGGCTCCAGCGTTCCTCCAT	666
QY	679	TTAATGGATCCAAAGTTCACAAGAAATACCTGTGATGGAATTTGATTTGGCAAGTTTG	738
Db	667	TTAATGGATCCAAAGTTCACAAGAAATACCTGTGATGGAATTTGATTTGGCAAGTTTG	726
QY	739	GAGGAGCTTGAACCTCACCATGACCAAGTTCATGATTTGTGCATCCTGTGTGATGTAC	798
Db	727	GAGGAGCTTGAACCTCACCATGACCAAGTTCATGATTTGTGCATCCTGTGTGATGTAC	786
QY	729	TATTTGTATACATCAAAAGGTATGGGGGGGCAAAACACTCTGGAACCTTATTCGCAACAT	858
Db	787	TATTTGTATACATCAAAAGGTATGGGGGGGCAAAACACTCTGGAACCTTATTCGCAACAT	846
QY	859	GGGTCCATAGAAAGCATCTTGAGAAATCTTAATAGAAGATATCAAAATTCCTGAGAC	918
Db	847	GGGTCCATAGAAAGCATCTTGAGAAATCTTAATAGAAGATATCAAAATTCCTGAGAC	906
QY	919	TGGCTTACCAAGAAAGCTCGACGCTTGTTCAGAGAGCCTAATGTCACTTTGGATATTCT	978
Db	907	TGGCTTACCAAGAAAGCTCGACGCTTGTTCAGAGAGCCTAATGTCACTTTGGATATTCT	966
QY	979	GAGCTAAATATGACTGCGACCTGATGAGGAGGCTCATAGTTTCCGTGTAAGAAATAT	103
Db	967	GAGCTAAATATGACTGCGACCTGATGAGGAGGCTCATAGTTTCCGTGTAAGAAATAT	102
QY	1039	GGTTCAATGAAGATCGGGGTACAAAGGCCATAGAGAAGATCAAAATTCGCAAGATATA	109
Db	1027	GGTTCAATGAAGATCGGGGTACAAAGGCCATAGAGAAGATCAAAATTCGCAAGATATA	108
QY	1099	TGCTGCGAAGGAAGACTCGAGTCTTTTTCAGGCCAATGCGCACCATATAGACACGCTA	115
Db	1087	TGCTGCGAAGGAAGACTCGAGTCTTTTTCAGGCCAATGCGCACCATATAGACACGCTA	114
QY	1159	AAACGGAAGGAGACTCTCGATATAAACAAGCAAGGACACTCGGAACAGAAACAAAGGCT	121
Db	1147	AAACGGAAGGAGACTCTCGATATAAACAAGCAAGGACACTCGGAACAGAAACAAAGGCT	120
QY	1219	GGTGAAGGAAGAAATATCTTTGATCTGTGATGTACAACTAGCACTACGAAAGCAGCGG	127
Db	1207	GGTGAAGGAAGAAATATCTTTGATCTGTGATGTACAACTAGCACTACGAAAGCAGCGG	126
QY	1279	TGGCATGATTCACCTTGGCTAGATTTATTTAACTCCCTGTTTAACTAGACCTTTGGTGA	133
Db	1267	TGGCATGATTCACCTTGGCTAGATTTATTTAACTCCCTGTTTAACTAGACCTTTGGTGA	132
QY	1339	AGTTTGCCCATGTTTCAAGCTGGGGGTAGTAGTTGTGTGAAGACATTTGGTGCAC	139
Db	1327	AGTTTGCCCATGTTTCAAGCTGGGGGTAGTAGTTGTGTGAAGACATTTGGTGCAC	138
QY	1399	GTAACAAAACCTTATCGCTGTTTTTAACTCTGTCTTTGAAGATGATATGCCAGTAA	145
Db	1387	GTAACAAAACCTTATCGCTGTTTTTAACTCTGTCTTTGAAGATGATATGCCAGTAA	144
QY	1459	AAAAAAAAAAAAAAAAAAAA 1475	
Db	1447	AAAAAAAAAAAAAAAAAAAA 1463	
RESULT 3			
US-09-426-557-3			
Sequence 3, Application US/09426557			
Patent No. 6232527			
GENERAL INFORMATION:			
APPLICANT: Mahajan, Pramod B.			
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses			
FILE REFERENCE: 0961			
CURRENT APPLICATION NUMBER: US/09/426,557			
CURRENT FILING DATE: 1999-10-22			

```

? EARLIER APPLICATION NUMBER: 60/112,332
? EARLIER FILING DATE: 1998-12-15
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FastSeq for Windows Version 3.0.
? SEQ ID NO 3
? LENGTH: 1541
? TYPE: DNA
? ORGANISM: Zea mays
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (79)...(1215)
? OS-09-426-557-3

```

Query Match	Score	DB 4;	Length
94.88;	1401.2;		1541
Post Local Similarity	00.19;		
Post Local Similarity	00.19;		

Matches 1409; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY	33	GGGGTTTCTTCCGGCAGCTCCGGCTCAGCGCCGCCGACACCGGCACAGCGCCGAG	92
Db	15	GGGGTTTCTTCCGGCAGCTCCGGCTCAGCGCCGCCGACACCGGCACAGCGCCGAG	74
QY	93	CGAGATGGGCATCAAGGGTTTGAGAAACTGCTGGCGGACATGGCGCCCAAGCGGATGAA	152
Db	75	CGAGATGGGCATCAAGGGTTTGAGAAACTGCTGGCGGACATGGCGCCCAAGCGGATGAA	134
QY	153	GGAGAGAGATTTCCAGAGCTACTTTCGGCCGAAATATGCGCTCCAGCGCAGATGACAT	212
Db	135	GGAGAGAGATTTCCAGAGCTACTTTCGGCCGAAATATGCGCTCCAGCGCAGATGACAT	194
QY	213	CTACCACTTCTTGATAGTACTTGGAAAGCAGCGCATGGAACCTCTCAAAATGAAGCTGG	272
Db	195	ATACCACTTCTTGATAGTACTTGGAAAGCAGCGCATGGAACCTCTCAAAATGAAGCTGG	255
QY	273	TGAAGTCACTAGTCATTTGCAAGSAATGTTCAACGGACATAAGATTACTGGAAGCGG	333
Db	255	TGAAGTCACTAGTCATTTGCAAGSAATGTTCAACGGACATAAGATTACTGGAAGCGG	314
QY	333	AATCAAGCACTTTATGTTTGGATGGCAACCCCTCGATTTGSAAGAACAAAGACTTCG	392
Db	315	AATCAAGCACTTTATGTTTGGATGGCAACCCCTCGATTTGSAAGAACAAAGACTTCG	374
QY	393	TAAAGATACTCAAAAGAGATGATGCAACCAAGATCTGACTAGGACAGTAGAGTAGG	452
Db	375	TAAAGATACTCAAAAGAGATGATGCAACCAAGATCTGACTAGGACAGTAGAGTAGG	433
QY	453	AGATTAAGATCCGATTGTAATAATTGAGCAGAGAGACTGTAAAGTCAACAGCAACAA	512
Db	435	AGATTAAGATCCGATTGTAATAATTGAGCAGAGAGACTGTAAAGTCAACAGCAACAA	494
QY	513	CGAAGATTGTAAAGGACATTTAAAGCTTATGGGGGTTCCTGTTAAGAGCACCTTCGA	572
Db	495	CGAAGATTGTAAAGGACATTTAAAGCTTATGGGGGTTCCTGTTAAGAGCACCTTCGA	554
QY	573	AGCAGAGCAGAAATGTCCAGCCCTTTTGCATTAACGATAAAGTGTTCGCTTTCCTTCA	632
Db	555	AGCAGAGCAGAAATGTCCAGCCCTTTTGCATTAACGATAAAGTGTTCGCTTTCCTTCA	614
QY	633	AGATTAAGGACTCCCTTACTTTTGGGGCTCCAGCGTTCCTTCGTCAATTAATGATCCAG	692
Db	615	AGATTAAGGACTCCCTTACTTTTGGGGCTCCAGCGTTCCTTCGTCAATTAATGATCCAG	674
QY	693	TTTCAAGAAATTAACCTTGATGAGATTTGATGTTGCCAAGTTTGGAGAGGCTTGACT	752
Db	675	TTTCAAGAAATTAACCTTGATGAGATTTGATGTTGCCAAGTTTGGAGAGGCTTGACT	734
QY	753	CACCATGAGCAGTTTCATTTGATTTGTGCATTCCTGTGTGATATGACATTTGTATAGCAT	812
Db	735	CACCATGAGCAGTTTCATTTGATTTGTGCATTCCTGTGTGATATGACATTTGTATAGCAT	794
QY	813	CAAAGGTATCGGGGGGCAACAGCTCGAAGCAATTAATTCGTCACATGGGTCATAGAAAG	872
Db	795	CAAAGGTATCGGGGGGCAACAGCTCGAAGCAATTAATTCGTCACATGGGTCATAGAAAG	854

QY	873	CATTGGAGAACTCTTAATTAAGACAGATATCAAAATTCCTGGAGGACTGGCCCTTACCAGA	932
Db	855	CATTGGAGAACTCTTAATTAAGACAGATATCAAAATTCCTGGAGGACTGGCCCTTACCAGA	914
QY	933	AGCTCGAGCGCTGTGTTCAAGAGAGCCCTAATGCATTTTGATTTCCGTGAGCTTAAATGAC	992
Db	915	AGCTCGAGCGCTGTGTTCAAGAGAGCCCTAATGCATTTTGATTTCCGTGAGCTTAAATGAC	974
QY	993	TGCACCTGATGAGGAGGGCTCATTAAGTTTCTGTGTAAGAATTAATGCTTCATGACGA	1052
Db	975	TGCACCTGATGAGGAGGGCTCATTAAGTTTCTGTGTAAGAATTAATGCTTCACGAGGA	1034
QY	1053	TGCGGTCACAAAGGCCCTATAGAGAAATCAAAATGTCCAGAAATTAATGTCTGCACAGSAG	1112
Db	1035	TGCGGTCAGAAAGGCCCTATAGAGAAATCAAAATGTCCAGAAATTAATGTCTGCCAAGSAG	1094
QY	1113	ACTGAGATCTCTTTTCAAGCAACATCGACACATCAGCACGCTTAAACGGAGAGAC	1172
Db	1095	ACTGAGATCTCTTTTCAAGCAACATCGACACATCAGCACCCCTTAAACGGAGAGAC	1154
QY	1173	TTTGGGATPAAAAACAAGCAGGCTGCGAACAAGAAACAAGGCTGTGGAAGAAGA	1232
Db	1155	TTTGGGATPAAAAACAAGCAGGCTGCGAACAAGAAACAAGGCTGTGGAAGAAGA	1214
QY	1233	ATAATCTTGATGCTTATATTAACAACACACACACAGCAAGCGGTGGCATGATCACTT	1292
Db	1215	ATAATCTTGATGCTTATATTAACAACACACACACAGCAAGCGGTGGCGATGATCACTT	1274
QY	1293	CGCCTAGATTATTAACCTCCCTGTTTTAACTCAGACCTTTGGGAAGATTTCCTCATGTT	1352
Db	1275	CGCCTAGATTATTAACCTCCCTGTTTTAACTCAGAGCTTTGGTAAAGTTTCTCATGTT	1334
QY	1353	TCAAGCTGGGGTAAGTTAGTTGTGTTGAAGAGATTGTTACCAAGTAACAAAACCTTAT	1412
Db	1335	TCAAGCTGGGGTAAGTTAGTTGTGTTGAAGAGATTGTTACCAAGTAACAAAACCTTAT	1394
QY	1413	CGCTGTTTTTACTCTTCTTGTCCTTTGAAGTATATATCCAGCT	1454
Db	1395	CGCTGTTTTTACTCTTCTTGTCCTTTGAAGTATATATCCAGCT	1436

```

RESULT 4
US-09-426-557-5
; Sequence 5, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(1173)
US-09-426-557-5

```

Query Match	90.3%	Score 1334.2	DB 4	Length 1381
Best Local Similarity	98.7%	Pred. No. 0		
Matches 1345; Conservative	0	Mismatches 18	Indels 0	Gaps 0

Oy	61 CGCGCCCGGCAACCCGGCAGAGCGCGGAGAGAGATGGGCATCAAGGGTTTGACGAAA	120
Db	1 CGACCACAGCGCTCCGGCCACAGCGCGCGGAGAGAGATGGGCATCAAGGGTTTGACGAAA	60

OY	121	CTTGGGGGACAAATGCGCCCAAGGGGATGAAGSACACAACTGGAGAGCTACTTGGC	180
Db	61	CTGTGGGGGACAAATGGGCCCAAGGGGATGAAGSACACAAAGTTGAGAGCTACTTGGC	120
OY	181	CGCAAAATCGCCGTCGACGGCCACATGAGCTATCCAGCTTCCGTATGATGTGAAG	240
Db	121	CGCAAAATCGCCGTCGACGGCCACATGAGCTATCCAGCTTCCGTATGATGTGAAG	180
OY	241	ACAGGCATGGAACCTCTCACAAATGAAGCTGTGAAGTCACTAATCATTTTGCAGGAATG	300
Db	181	ACAGGCATGGAACCTCTCACAAATGAAGCTGTGAAGTCACTAATCATTTTGCAGGAATG	240
OY	301	TTTCAACGGGCAATAAATTAATTAAGTGAAGGGGAATCAACCAAGTTATGTTTATATGC	360
Db	241	TTTCAACGGGCAATAAATTAATTAAGTGAAGGGGAATCAACCAAGTTATGTTTATATGC	300
OY	361	AAGCCTCCTGATATGAAGAAACAAGAACTTGCTTAAAGATTAAGTCTCAAAAAGAGATGATCA	420
Db	301	AAGCCTCCTGATATGAAGAAACAAGAACTTGCTTAAAGATTAAGTCTCAAAAAGAGATGATCA	360
OY	421	ACCAAAGATCTGACTGAGGCAATGAGAGTGAAGATTAAGATGCGATTGCAAAAATTGAGC	480
Db	361	ACCAAAGATCTGACTGAGGCAATGAGAGTGAAGATTAAGATGCGATTGCAAAAATTGAGC	420
OY	481	AAGGAGCTGTAAAGGTCCACAGGCAACACAAAGAAATTTGAACGATTAATTAAGCTT	540
Db	421	AAGGAGCTGTAAAGGTCCACAGGCAACACAAAGAAATTTGAACGATTAATTAAGCTT	480
OY	541	ATGGGGGTTTCTCTTTTGAAGAGGCACTTCTGAAGCAGAAAGATGTGACGCCCTTTGC	600
Db	481	ATGGGGGTTTCTCTTTTGAAGAGGCACTTCTGAAGCAGAAAGATGTGACGCCCTTTGC	540
OY	601	ATAAACGATTAAGGTGTTCGCTGTTCCTCAGAAGATTAAGACTCCCTTACTTTGGGCT	660
Db	541	ATAAACGATTAAGGTGTTCGCTGTTCCTCAGAAGATTAAGACTCCCTTACTTTGGGCT	600
OY	661	CCAGGTTCTTCTGCTCAATTAATGAGATCCAAAGTTCACAAAGAAATACCTGATGGAATTT	720
Db	601	CCAGGTTCTTCTGCTCAATTAATGAGATCCAAAGTTCACAAAGAAATACCTGATGGAATTT	660
OY	721	GATGTGGCAAGTTTTGGAGAGACTTGAACCTACCATGAGCACCTGTTGATTTGTGC	780
Db	661	GATGTGGCAAGTTTTGGAGAGACTTGAACCTACCATGAGCACCTGTTGATTTGTGC	720
OY	781	ATCCTGTGTGATGTGACTATTGTGATAGCATTAAGGATTCGGGGGCAACACCTGTG	840
Db	721	ATCCTGTGTGATGTGACTATTGTGATAGCATTAAGGATTCGGGGGCAACACCTGTG	780
OY	841	AAACTTATTCGTCAACATGGGTCCATGAAGAAGATCTTGAAGATTTTAATAAGACGA	900
Db	781	AAACTTATTCGTCAACATGGGTCCATGAAGAAGATCTTGAAGATTTTAATAAGACGA	840
OY	901	TATCAAAATTCCTGAGGAGCTGGCCTTACCAGAAGCTGCAGCTTGTTCAAAGAGCCTAAT	960
Db	841	TATCAAAATTCCTGAGGAGCTGGCCTTACCAGAAGCTGCAGCTTGTTCAAAGAGCCTAAT	900
OY	961	GTCACATTGGATTTCTGAGACTTAAATGAGCTGACCTGATAGGAGGGTCTCATAGT	1020
Db	901	GTCACATTGGATTTCTGAGACTTAAATGAGCTGACCTGATAGGAGGGTCTCATAGT	960
OY	1021	TTTCTCGTGAAGAAAGTATGTTTCATAGAGATTCGGGTGACAAAGGCCATTAAGAAATG	1080
Db	961	TTTCTCGTGAAGAAAGTATGTTTCATAGAGATTCGGGTGACAAAGGCCATTAAGAAATG	1020
OY	1081	AAATTCGCAGAAATTAATGCTGCCAAGAAAGCTGCAGCTCTTTTCAAGCAACTGCG	1140
Db	1021	AAATTCGCAGAAATTAATGCTGCCAAGAAAGCTGCAGCTCTTTTCAAGCAACTGCG	1080
OY	1141	ACCAATCAGCACCGCTAAAAACGAAAGAGACTTCGATAAAAACAAGCAAGCAGCTGGG	1200
Db	1081	ACCAATCAGCACCGCTAAAAACGAAAGAGACTTCGATAAAAACAAGCAAGCAGCTGGG	1140
OY	1201	AACAAGAAACAAGGCTGTGTGAAGAGAAATTAATCTTGGATGCTTATGATACACTTA	1260

|||||
Db 1141 AACAGAAAAGAGGCTGGGAGAAAGAAATATCTGGATGCTGATGACAACTA 1200
QY 1261 CGACTAGAAAAGCAGCGGTCATGATCATCTGCGCTAGATTATTTAACTCCCGTTTAA 1320
Db 1201 CGACTAGAAAAGCAGCGGTCGCGGTCATGATCATCTGCGCTAGATTATTTAACTCCCGTTTAA 1260
QY 1321 ACTGAGACCTTTGGTGAAGTTTGGCCCATGTTTCAGCTGGGGTAACTTAACTGTTG 1380
Db 1261 ACTGAGACCTTTGGTGAAGTTTGGCCCATGTTTCAGCTGGGGTAACTTAACTGTTG 1320
QY 1381 AAGGATTGGTGTACCAAGTACAAACTTATCGCTGTTT 1423
Db 1321 AAGGATTGGTGTACCAAGTACAAACTTATCGCTGTTT 1363

RESULT 5
US-08-455-968E-2
; Sequence 2, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-455-968E-2

Query Match 20.3%; Score 299.4; DB 2; Length 1144;
Best local Similarity 56.5%; Pred. No. 4.6e-73;
Matches 599; Conservative 0; Mismatches 456; Indels 6; Gaps 2;

QY 97 ATGGGATCAAGGCTTGGAGAAATGCTGCGGACATCGGCCAAGCGATGAAGAG 156
Db 1 ATGGGAATTCAGAGCCGCGCAAACTAATGCTGATGTGCCCCCACTGCGATCCGGGAG 60
QY 157 CACAAGTTCAGAGCTACTTCTGCGCCCAAAATGCGCGTCGACGCCACCATGAGCATCTAC 216
Db 61 AATGACATCAAGAGCTACTTCTGCGCCCAAAATGCGCGTCGATGATGAGCATTTAT 120
QY 217 CAGTTCCTGATAGTGTGAGAGACAGCATGGAACCTCACAATGAAGCTGAGGAA 276
Db 121 CAGTTCCTGATGCTGTTCCCGAG--GGTGGGATGCTGCGAGAAATGAGGAGGTGAG 177

QY 277 GTCACTAGTCATTGGCAAGAAATGTTCAACCGGACAAATTAAGATTACTGGAAGCGGAATC 336
Db 178 ACCACACAGCCACCTGATGGGATGTTCTACCGCACCATTTGCGATGATGAACAGGCAATC 237
QY 337 AAGCCAGTTATGTTTTTGGATGGAAGCCCTCTGATATGAAGAAACAAAGAACTTCTTAA 396
Db 238 AAGCCAGTTATGTTTTTGGATGGAAGCCCTCTGATATGAAGAAACAAAGAACTTCTTAA 297
QY 397 AGATATCTCAAAAAGATGATGCAACCAAGATGCTGACTGAGGCAATGAGTGAAGAT 456
Db 298 CGCAGTGAAGCGCGGCTGAGAGAGAGCAAGCAGCTGACAGCGCTCAAGCTGCTGGGGCC 357
QY 457 AAAGATGCGATTGAAAATTTAGAGCAAGAGACTGTAATTAATTAATTAATTAATTAATTA 516
Db 358 GAGCAGAGAGTGGAAAAATTTACTTAAGCGGCTGTGTAAGTCTACAGCAGCAGCATATAT 417
QY 517 GATTGTAAAGCACTATTAAAGACTTATGAGGCTTCTGTTGTAAGGCACTTCTTAAGCA 576
Db 418 GAGTGAACAACTATGCTGAGCCTCATGGCAATCCCTTATCTGATGCAACCCAGTAGAGCA 477
QY 577 GAACAGAAATGTGACGCCCTTGCATTAAGCATTAAGCTGTGCTGTTCAGAGAT 636
Db 478 GAGGCACAGCTGTGCTGCGCTGTGTAAGGCTGCAAAAGTCTATGCTGCGCTACCGAGAGAC 537
QY 637 AAGCACTCCCTTACTTTTGGGGCTCAAGGCTTCTGCTGCTTAATTAATTAATTAATTAAT 696
Db 538 ATGACATGCTCCTACCTTCGAGCGCTGTGTCTTAATGACACCTGACTGCAAGTGAAGCC 597
QY 697 AAGAAATACCTGTGATGAATTTGATGTTGCAAGGTTTGGAGAGCTTGAATCACTCAC 756
Db 598 AAAAAGCTGCAATTCAGGAATTCACCTGAGCGGATTTGAGAGAGCTGGGCTGAGAC 657
QY 757 ATGACACATGTCATGATTTTGTGATCTGTGTGATGTGCTATTTGATAGATCAAA 816
Db 658 CAGGAACAGTTTGTGATCTGTGATCTGTGATGATGATGATGATGATGATGATGATG 717
QY 817 GGTATCGGGGGGCAAAACAGCTCTGAACCTTATTCGTCACATGGGTGTCATGAAGAGCAT 876
Db 718 GGTATTTGGGCCCAAGCGGCTGTGACCTCTATCCAGAAAGCAAGAGATGAGAGATC 777
QY 877 TTGAGAAATCTTAATTAAGACAGATATCAAAATTCCTGAGAGCTGCGCTTACCAAGAACT 936
Db 778 GTGCGGAGACTTGAACCCCAACATACCTGTGCGAGAAATTTGCTGCCAAGAGAGCT 837
QY 937 CGAGCTTTGTCAGAGAGCTTAATG---TACTTTGGATATTTCTGAGCTTAATGAGCT 993
Db 838 CACCAGCTCTTCTTGAACCTGAGAGTGTGAGCCAGAGTCTGTGAGAGCTGAAAGTGGAGC 897
QY 994 GCACCTGATGAGAGGCTCATTAATGTTCTGTGTAAGATTAATGTTTCAATGAAGAT 1053
Db 898 GAGCAATTAAGAAAGAGCTGATCAATGTTCTATGTGTGAAAAGCAAGTTCTTGTAGAGAG 957
QY 1054 CGGCTGACAAAGGCCATAGAGAAATCAATCTGCCAAGAAATTAATCTGCGAAGAGAGA 1113
Db 958 CGAATTCGAGTGGGTCGAGAGAGCTGAGTAAGAGCCGCCAAGGACAGACCCAGAGGCCG 1017
QY 1114 CTGAGATCTTTTTCAGGCAACTGCCACCATACAGCACC 1154
Db 1018 CTGAGATCTTTTTCAGGTCAGGTCACCGGCTCACTCTCTTACAG 1058

RESULT 6
US-08-455-968E-28
; Sequence 28, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:

Query Match	20.3%	Score 299.4	DB 2	Length 1144
Best Local Similarity	50.5%	Pred. No. 4.6e-73		
Matches 599, Conservative	0	Mismatches 456	Indels 6	Gaps 2

```

Db 538 ATGACATGCCCTACACCTTCGGGAGGCCCTGTGCTAATGGAGACACTGACTGACAGTGAACC 597
Oy 697 AAGAAATACCTGTGATGAGAAATTTGATGTTGCCAAGSTTTTGGAGAGAGCTTGAACTCACC 756
Db 598 AAAAAGCTGCCAATCCGAGAAATTCACCTAGCCGGGATTCGTGGAGAGCTGGGGCTTGAAC 657
Oy 757 ATGAGACCAATTCATGATTTGTGCATCCTGTGTGATGTGACTATCTATGTATAGATCAAA 816
Db 658 CAGAGACAGATTTGTGATCTGTGTGATCTCTGTAGGCAAGTCACTACTGTAGAGATTCGG 717
Oy 817 GGTATGGGGGGGCAACAGCTCTGAAACTTATTTGTCACATGGGTTCATAGAACATC 876
Db 718 GGTATTTGGGCCCAAGCGGGCTGTGCACTCATCCAGAAAGCAAGAGCATCGAGGATC 777
Oy 877 TTGGGAATCTTAATTAAGACAGATATCAAAATTCGTAGGACCTGGGCTTATCCAAAGCT 936
Db 778 GTGGGGCACTTTGACCCCAACAGATACCTCTGTGCCAGAAATTTGGCTCCACAAAGAGCT 837
Oy 937 CGAGCGTGTTCGAAGGAGGCTTAATG--TCACTTTGGATATTCCTGAGCTAAATGGACT 993
Db 838 CACCAAGCTCTTCTTGGAAACCTGAGAGTCTGTGACCAGACAGATCTGTGGAGCTGAAGTGAAC 897
Oy 994 GCACCTGATGAGAGGGTCTCATAGTTTCTGTGTAAAGATATATGTTTCATGTAAGT 105
Db 898 GAGCCAAATGGAAGAGAGCTGTATCAAGTTTATGTGTGTAAGCAAGCATGTTCTTGAGAG 957
Oy 1054 CGCGTGACAAAGGCCATTAAGAAATCAAAATCTGCCAAGATAATCTGTCGAAGAGAA 111
Db 958 CGAATCCCGCATGGGGGTTCAGAGGCTGATGAAGAGCCGCCAAGGACAGACCCAGGGCCG 101
Oy 1114 CTCGAGTCCCTTTTCAAGCCACTGCCACACATCAGCACC 1154
Db 1018 CTGATGATTTCTTCAAGGTGACCGGCTCACTCTTTCAGC 1058

RESULT 7
US-08-455-968E-9
; Sequence 9, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Jin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18995-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

1  RESULT 7
2  US-08-455-968E-9
3  ; Sequence 9, Application US/08455968E
4  ; Patent No. 5874283
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Harrington, John L.
7  ; APPLICANT: Hsieh, Chih-Lin
8  ; APPLICANT: Lieber, Michael
9  ; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
10 ; NUMBER OF SEQUENCES: 63
11 ; CORRESPONDENCE ADDRESS:
12 ; ADDRESSEE: Townsend and Townsend and Crew LLP
13 ; STREET: Two Embarcadero Center, 8th Floor
14 ; CITY: San Francisco
15 ; STATE: California
16 ; COUNTRY: USA
17 ; ZIP: 94111-3834
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Floppy disk
20 ; COMPUTER: IBM PC compatible
21 ; OPERATING SYSTEM: PC-DOS/MS-DOS
22 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
23 ; CURRENT APPLICATION DATA:
24 ; APPLICATION NUMBER: US/08/455,968E
25 ; FILING DATE: 30-MAY-1995
26 ; CLASSIFICATION: 435
27 ; ATTORNEY/AGENT INFORMATION:
28 ; NAME: Apple, Randolph T.
29 ; REGISTRATION NUMBER: 36,429
30 ; REFERENCE/DOCKET NUMBER: 18995-000100
31 ; TELECOMMUNICATION INFORMATION:
32 ; TELEPHONE: 415-576-0200
33 ; TELEFAX: 415-576-0300
34 ; INFORMATION FOR SEQ ID NO: 9:
35 ; SEQUENCE CHARACTERISTICS:
36 ; LENGTH: 2033 base pairs
37 ; TYPE: nucleic acid
38 ; STRANDEDNESS: single

```



```

:
: MOLECULE TYPE: DNA (genomic)
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 104..1237
US-08-455-968E-9

Query Match      18.6%; Score 275; DB 2; Length 2033;
Best Local Similarity 55.8%; Pred. No. 3,3e-66;
Matches 637; Conservative 0; Mismatches 480; Indels 24; Gaps 5;

QY 97 ATGGGCAATCAAGGGTTTGAGCAAACTGCTGGCGGACATATGGGCCCAAGCCATGAGAG 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 104 ATGAAATTCACGGCCTTGCCAAACTAATCTGATGTGCGCCCAAGTCATCCGTAG 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 CAGAAGTTGAGAGCTTACTTGGCCGCAAAATCGCCGTGCGCAGCAGCATGAGCTTAC 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 AATGACATCAAGACTACTTTGGTCGCAAAAGTGGCCATCGATGCTCATGACATCTAC 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 CAGTTCTGATAGTAGTTGGAGAGCAGCGCATGCAACTCTCACAATGAAGCTGGTGA 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 CAGTTCTGATAGTAGTTGGTCGCAAG--GTTGGGATGTGTGCAAGAACGAGAGGTTAG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 GTCACATGCTATTGTCAGAGAAATGTTCAACCGGACAAATAGATTACTGAGAGCGGAA 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 ACCACACGCC---TGATGGGCAATGTTTACCCGTACCATGCG---CATGGAAATGGCATC 334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 337 AAGCCAGTTATGTTTTTGTATGAGCGACCTCTGATATGAAGAAACAGAACTTGCTAA 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 AAGCCTGTGCTGCTTTGATGTCAAACACACAGCTGAAGTGAAGGAGACTGGCCAG 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 397 AGATACCAAAAAGATGATGACCAACCAAGATCTGACTGAGCGAGTAGAGCTAGAGAT 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 CGCAGTAGAGGCGCGCGAGGCTGAGAGAGCACTGCAAGCAAGCTCAAGCAGCTGGGATG 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 AAGATGCGATTGAAATATTGAGCAAGAGGACCTGTAAGAGTCACAAAGGCAACACAGAA 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 GAGAGAGAGGTGGAGAGTTCCACCAAGAGGCTGTGAAGGTCAACCAACACACAAATGAT 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 GATTGTAACGACTATTAAAGACTTATGGGGTCTCTGTTAGAGGACCTTTCGAAGA 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 GAGTGCAAAACACCTGCTGAGCGCTCATGGGACATCCCTTACCTGTGATGACCCAGGAG 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 577 GAAGCAAAATGTCAGGCGCTTTCATTAACGATTAAGGTGTTGCTGTTCGTTCAAGAT 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 GAGGCCAGCTGTGCTGCTGGCAAGAGCTGGCAAGCTAATGCTGGCGCCACGAGAGAC 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 637 AAGGACTCCCTTACTTTTGGGGGCTCCACGGTTCCTGTCATTTAATGATCCAAGTTCC 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 ATGGATGCTCACTTTTGGCAGGCCGCTGTATGACACTTAACCTGCCAGTGAAGCC 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 697 AAGAAATACCTGTGATGAATTTGATGTTGCCAAGGTTTTTGAAGAGCTTGAACAC 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 695 AAGAAAGCTGCCATCCAAAGATTCATCTAGACCGCTCCCTGCGAGAGCTGGGCTGAAC 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 757 ATGGACCAATTCATTTGTTGTGATCTCTGTTGATGATGATGATGATGATGATGATG 816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 755 CAGGAGAGATTTGTGATCTGTGATCTCTGTGGTGAAGCACTGCGAGAGATCCGT 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 GGTATCGGGGGGCAAAACAGCTCTGAACCTTATTCGTCACATGAGGTCATAGAAAGCAT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 815 GGCATTTGGCCCAAGCGGGCTGTGGATCTCATCGAAGAACTAAGACATCGAGGAATC 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 TTGGAGATCTTATTAAGACAGATATCAAAATTCCTGAGAGCTGGCTTACCAAGAAGCT 936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 875 GTAGAGGGGTGAGACCCACCAAGTACCCCGTTCCAGAGAACTGGCTCCACAAAGAGCC 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 CGAGGCTTTGTCAGAGGCTAA---TGTACATTTGATATTTCTCGTGAATTAATGACT 993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 935 CAGCAGCTCTTCTCGAGGAGAGAGTAGTAGACCCAGAGCTGTGTGAGCTGAAGTGAAGC 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 994 GCACTGTAGTAGAGAGGCTCTCATTAAGTTTCTGTGTAAGATATGATTAATGAAGAT 1053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

Db 995 GAGCCAAATGAAGAGAGTTGCTCAAAATTTATGTGTGTAAGAAACAGTTTTTTGAAG 1054
QY 1054 CGGGTGACAAAGGCCATAGAGATGATCAATCTGCCAGAAATTAATTCGTGCAAGAGAG 1113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 CGAATTCGACGTGGGGGTCAAGGGGCTGAGTAGAAGACCCCGCAGGACGACCCAGGAG 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1114 CTCGAGTCCTTTTTCAGCCCACTGCTCCACACATCAGACCGCTAATAACGAGAGACT 1173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 CTCGATGATTTCTTCAGAGTGACAGGCTCAGCTCTCTCAGC-----TAAAGCC 1162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1174 TCGATTAACAGACAGACGAGCTGCGAACAGAAACAAAGCGCTGTGGAAAGAGAAA 1233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1163 AAGGAGCCAGAACCCAGGCGCTGCTAAGAAAGAACAGACTGGGGAGAGCGGGAAG 1222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1234 T 1234
Db 1223 T 1223
```

```

RESULT 8
US-08-455-968E-4
: Sequence 4, Application US/08455968E
: Patent No. 5874283
:
: GENERAL INFORMATION:
: APPLICANT: Harrington, John L.
: APPLICANT: Hsieh, Chih-Lin
: APPLICANT: Lieber, Michael
: TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
: NUMBER OF SEQUENCES: 63
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentLin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,968E
: FILING DATE: 30-MAY-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Apple, Randolph T.
: REGISTRATION NUMBER: 36,429
: REFERENCE/DOCKET NUMBER: 18985-000100
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-576-0200
: TELEFAX: 415-576-0300
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1930 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
US-08-455-968E-4

Query Match      18.3%; Score 270.6; DB 2; Length 1930;
Best Local Similarity 55.0%; Pred. No. 5.2e-65;
Matches 628; Conservative 0; Mismatches 489; Indels 24; Gaps 4;
```

```

QY 97 ATGGGCAATCAAGGGTTTGAGCAAACTGCTGGCGGACATATGGGCCCAAGCCATGAGAG 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGGGAAATTCACGGCCTTGCCAAACTAATTCGATGTGCGCCCAAGTCATCCGTAG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 CAGAAGTTGAGAGCTACTTTCGCGCCGCAAAATCGCGTCGACGCGCAGCATGAGCATCTAC 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AATGACATCAAGAGCTACTTTGTGCTGAAGTGGCCATCGATGCTCTCATGAGCATCTAC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY 217 CAGTTCGATAGTATGTTGGAAGGACGACATGGAACCTCTCACAAATGAAGTGTGAA 276
|||||
Db 121 CAGTTCGATAGTATGTTGCTGAC---GGTGGGATGTCTGCGAAGAGAGAGGATG 177
QY 277 GTCACATGATTTGCAAGGAATGTCAACCGACAAATAGATTACTGGAAGCGGGAAT 336
|||||
Db 178 ACCACAGCGCTGATGGGATGTAT-----GGCAACCATCCCATGAGCAATGGCATC 231
QY 337 AAGCCAGTTATGTTTGTGATGCAAGCCTCTCTATATGAAGAAACAAGAACTTCTTAA 396
|||||
Db 232 AAGCTGTATGCTCTGTGATGCAACCCACACAGCTGAAGTCAAGGAGCTTCCGCAAG 291
QY 397 AGATATCAAAAAGATGATGACAACCAAGATCTGACTGAGCGAGTAGAGTGGAGAT 456
292 CGCAGTGAAGAGCGCGCGGAGGCTGACAGCAACTGCACAGGCTCAGAGAGCTGGGATG 351
QY 457 AAGATGCGATTTGAAAAATTTGAGCAAGAGACTGTAAAGTTCACAAAGCAACAGCAA 516
352 GAGGAGGAGGTGGAAGATTACCAAGAGGCTCGTGAAGGTCACCAACAGCAACATGAT 411
QY 517 GATTGTAAACGACTATTAAAGCTATGGGGGTCTCTGTTAGAGGCACTTCTGAAGA 576
412 GAGTGAACACCTCTGTAGGCTCATGGGCTACCTTGATGACCCAGCGAGGCA 471
QY 577 GAAGCAGAAATGCGAGCCCTTTCATTAACGATAGAGTGTGCGCTGCTTCAGAAAT 636
472 GAGGCCAGCTTCTGCTGCTGCGCAAAAGCTGCGCAAAAGTCTATGCTCGCGCCACGAGAG 531
QY 637 AAGGACTCCTTACTTTTGGGGCTCCAGGTTCTCTGCTATTAATGATCCAAGTTCC 696
532 ATGAGCTCTCAGCTTTTGGCAGCCCGCTGCTAAATGCGACACTTAAGTCCAGTAGAGCC 591
QY 697 AAGAAAACTCTGTATGATGATTTGATGTTGCAAGTTTGGAGGACTTGAATCACC 756
592 AAGAGGCTGCGCATCAAGATTCATGTAGCGCGCTGCGACGAGCTGGTCTGTAC 651
QY 757 ATGACCACTTATGATTTGATGATCCTGTGTGATGATGATGATGATGATGATGATGAT 816
652 CAGGAGCAGTTTGTGATCTGTGATCTGCTGCTGGGTAGCAGTACTGGAGAGCATCCGT 711
QY 817 GGTATCGGGGGGCAACAGCTCTGAACCTTATGTCACATGGTCCATGAAGAACATC 876
712 GGCATTTGGCGCAAGCGGGCTGTGATCTCATCCAGAAACATTAAGCATGAGAGATC 771
QY 877 TTGGAGATTTTAATTAAGACAGATATCAAAATCTCTGAGGACTGGCTTACCAAGAGCT 936
772 GTGAGGCGGCTGGACCCGACAGATACCCCGTTCCAGAGAACTGGCTCCACAGAGAGCC 831
QY 937 CGAGCTTGTTCAGAGAGCTTAA--TGTCACTTTGATATTTCTGAGCTTAAATGACT 993
832 CAGCAGCTTCTTCTGAGAGCCAGAAAGTACTGAGCCAGAGTGTGTGAGCTGAAAGTGAGC 891
QY 994 GCACCTGTATGAGAGAGGTCTCATTAAGTTTCTGTGTAAGATATGTTTCAATTAAT 1053
892 GAGCCAAATGAAGAAAGATTTGTCAAAATTTATGTGTGTAAGAACAGTTTCTTAAAG 951
QY 1054 CGGGTGAACAAGGCGCATGAGAAATCAATCTGCCAAGAAATTAATTCGCGCAAGAGAA 1113
952 CGAATTCGACAGTGGGCTGAGAGGCTGAGTAAGAGCCGCCAGGAGGAGCAACCCAGGAGCC 1011
Db 1012 CTCGATGATTTCTTCAAGGTGACAGGCTCAGCTCTCTCAGC-----TAAAGGC 1059
QY 1114 CTCGATCCTTTTTCAGCCCACTGCCACACATCAGCAGCAGCTTAAAGCGAAGAGACT 1173
1114 TCGGATTAACAAGCAGGAGCTGCGAACAAGAAACAAGGCTGTGTGAAAGAGAA 1233
Db 1060 AAGGAGCCAGAAACCAAGGGGCTGTCTAAGAAAGAAAGAAAGACTGTGGGAGCGGGAAG 1119
QY 1234 T 1234
Db 1120 T 1120

RESULT 9
US-08-455-968E-6
; Sequence 6, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-455-968E-6
Query Match 18.0%; Score 265.4; DB 2; Length 1149;
Best Local Similarity 55.2%; Pred. No. 1,1e-63;
Matches 596; Conservative 0; Mismatches 456; Indels 27; Gaps 3;
QY 97 ATGGCATCAAGGCTTTACCAAACTGCTGGGCGACATGGGCGCCCAAGCGCATGAGAG 156
Db 1 ATGGGATTTAAAGGTTTGAATGCAATTAATTCGGAACATGTTCCCTGCTATCAGGAAA 60
QY 157 CAGAGTTCGAGAGCTTCTCGCGCGCAAAATCGCCGTGACGCGCAGCATGAGATCTAC 216
Db 61 AGCGATATCAAGAGCTTTTTCGCGCAAGAAAGTTGCGCATGATGCTATATCTATAT 120
QY 217 CAGTTCGATAGTATGTTGGAAGGACGACATGAACTCTCAAAATGAAGCTGTGAA 276
Db 121 CAGTTTAAATGCTGTAAACACAGACAGACGCTGGCGAGTTGACCAATGAAGCCGCTGA 180
QY 277 GTCATAGTCAATTCGCAAGAAATGTTCAACCGGACAAATTAAGTTTCTGGAAGCGGGAATC 336
Db 181 ACAAGTCACTGATGAGGTATGTTTATAGACACAGTGAATGATGATGATGATGATGATGAT 240
QY 337 AAGCAGTTTATGTTTGTGATGCGCAAGGCTCTGTATGAGAAACAAGAACTTGTCTAA 396
Db 241 AAGCTTGTATATGTTTGTGACGCGCAAACTCTCAGCTTTGAATTCATGAGTGTGCAAG 300
QY 397 AGATATCAAAAAGATGATGACAACCAAGATCTGACTGAGGCAAGTGAAGTGAAGAT 456
Db 301 CGGCTTCAAGAGAGGTGGAACAGAAAAAAGTGCAGAGGCAACAAGAAATTTGAA 360
QY 457 AAGATGCGATTTGAAAAATTTGAGCAAGAGCACTGTAAAGTTCACAAAGCAACAGCAAG 516
Db 361 AAGATGAAGCAAGAAAG-----AGATTGTGAAGTCTCAAAAGAGCATATATGA 411

```

QY 517 GATTGTAAGACATATTAGACTTATTGGGGTTCCTTGTGTAGAGGACCTTCTGAGCA 576
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 412 GAAGCCCAAAATTAAGACTATAGGACTATAGGAAATCCCATATATATATACGGCAAGAGCT 471
QY 577 GAGCAGAAATGTGAGCCCTTTGCTTAAGCATAGAGTATGCTCTTCTCAGAAAT 636
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 472 GAGGCTCATATGCTGAGTTGGCAAGAGGAAAGAGTATGCTCCGACAGTGAAGAT 531
QY 637 AAGGACTCCCTACTTTTGGGGCTCCAGCGTTCCTTCCTGATTTAATGATCCAGTTCC 696
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 532 ATGGACACACTCTGTATTAGAACCCCTTCTTGTGAGACATTTGACTTTTTCAGAGGC 591
QY 697 AAGAAATACCTGTGATGATTTGATGTCAGAGTGTGGAGAGCTTGAACCTACCC 756
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 592 AAGAAGAACCGATTACAGAAATAGATAGCATTTAGTTTGTAGAGACTCCGACTTGACA 651
QY 757 ATGGACAGTTGATTTGATTTGCTCTCTGTGATGTGATGATGATGATGATGATGAT 816
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 652 ATAGAGCAGTTTGTGATCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 711
QY 817 GGTATCGGGGGGCAACACGCTCTGAACTTATTCGTCACATGGTCCATAGAAAGCATC 876
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 712 GGTGTGGTCCAGTGACAGCGCTTAATTTGATTAAGAACGATGATGATGATGATGATGAT 771
QY 877 TTGGGAATCTT-----ATAAGACGATATCAATTTCTGAGGACTGG 921
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 772 GTGGGTTTATTGAATCTGGGAGTCAACACACTTAATGAAATCCGAGAGACTGG 831
QY 922 CCTTACCAAGAGCTCGACGCTTGTTCAGAGCGCTTAATTCACCTTGTGNTA---TTCT 978
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 832 CCTTACCAAGAGAGAAATGCTGTTCTTGACCTCAAGTTATAGTGTGATGATGATGAT 891
QY 979 GAGCTAAATGAGTGCACCGCTGATGAGAGGCTGCTGATGATGATGATGATGATGATGAT 1038
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 892 AACTTGAATATGCTGCCACCAAGAGAGAGAACTTATGCAATTTATGATGATGATGATGAT 951
QY 1039 GGTTCATGAAGATCGGGTGACAAAGGCCATAGAGAATCAATCTGCCAGAGATAAA 1098
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 952 AATTTCAGTGAAGAAAGATTAACTGTGATATCAAGATTGAAAGAAAGCTTGAATCT 1011
QY 1099 TCGTCGAGAGAGAGTGCAGTCTTTTTCAGGCCAAGCTCCGACCATAGACAGCT 1157
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1012 GGCATTTCAGGAGTATGATGGTTCTTCCAAGTGTGCTTAAGACAAAGAGACACT 1070

RESULT 10
US-08-757-653-175
; Sequence 175, Application US/08757653
; Patent No. 5843669
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; TITLE OF INVENTION: Thermostable PEN-1 Endonucleases
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,653
; FILING DATE:
; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 175:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1020
; US-08-757-653-175

Query Match 11.0%; Score 162.8; DB 2; Length 1023;
Best Local Similarity 55.0%; Pred. No. 2.1e-35;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Caps 0;

QY 146 CGATTGAAGAGCAGAGTTGCGAGCTACTTGGCGCAAAATCGCGTGCAGCCAGCA 205
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 29 CAAGAAAGAAATTTAGATTTAGAAACCTATACGGGAAAAAATGCAATGACCTCTTA 88
QY 206 TGAGCATCTACCGTTCCTGATAGTATTGGAAGGACAGGCAATGCAATGCAATG 265
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 89 ATCAATCTCCCAATTTTGTCCACAAATPAGACAGAAAGATGGAATCTCACTTATGATT 148
QY 266 AAGTGCTGAGTCACTACTACTATTGCAAGAAATGTCACCGGACATATAGATTACTGG 325
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 149 CAAGAGGTAAATTAACCTCCACCTAAGGGGCTCTTTTACAGGACATATAACTATTG 208
QY 326 AAGCGGAATCAAGCCAGTTTATTTTGTATGCGCAAGCTCTGATATGAAGAAAG 385
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 209 AGCGTGAATTAACCTGTGTATGTTTGTATGAGGAACCTCCAAATTCAAAAAGAG 268
QY 386 AACTTGTAAATGATCTCAAAAAGAGATGATCAACCAAGATCTGAGGAGCTAG 445
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 269 AGCTCGAAAAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328
QY 446 AGGTAGAGTAAAGATGCGATTGGAATAATTGACCAAGAGACTTAAGTCAAGAGC 505
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 329 AAAAGGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
QY 506 AACCAAGAGAGATTTGTAACGACTATTAAAGACTTATGAGGGGTTCTGTTAGAGCAC 565
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 389 TGTCTATCGAGATGCAAAAAAACCCTTAGAGCTTATGGAATTCCTATGATCAAGCAC 448
QY 566 CTTCGAGAGAGAGAGAGATGTCAGCCCTTGTGATTAAGCATTAAGTATAGGTGCTG 625
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 449 CTAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
QY 626 CTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 509 CTAGTCAAGATTCAGATCCCTCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 568
QY 686 ATCCAGTTCAGAGAAATTAACCTGATGATGATGATGATGATGATGATGATGATGAT 727
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 569 TACACGAGAAAGAAAGAGTTCCTGGGAAAAATGCTACCTCG 610

RESULT 11
US-08-823-516-78
; Sequence 78, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	11.0%	Score 162.8	DB 2	Length 1023
Best Local Similarity	55.0%	Pred. No. 2.1e-35		
Matches	320	Conservative	0	Mismatches 262
				Indels 0
				Gaps 0
QY	146	CGATGAAGAGACGAAGTTCGAGAGCTACTTCGCGCCGCAAAATGCGCGGTGAGGCCAGCA	205	
Db	29	CAGAAAAAGAAATTTGATGTGAAAACCTTATGCGGAAAAAATATGCAATGCAGCCCTCTTA	88	
QY	206	TGACCATCTACCGAGTTCCTGATGTAGTACGTTGGAAAGACAGCGCATGGAAACCTGCAAAATG	265	
Db	89	ATGCAATCTACCAATTTTTGTGTCACCAATTAAGACAAGAAAGATGGAATCTCCACTTATGTGATT	148	
QY	266	AAGCTGTGAGAGTCACTAGTCAATTTTCACAGAGATTTTCACCGGCACATAAGATTAAGTTAGTG	325	
Db	149	CAAAAGGTAGAAATTAACCTCCCACTCAAGGCGCTCTTTTACGAGCAATAAACTTAATGG	208	
QY	326	AAGCGGGAATCAAGCCAGTTTATGTTTTGATGGCAAGCCTCTCGATATGGAAGAAACAG	385	

Db	209	AGGCTGAATAAACCCTGTGTATTTTGTGTGAGAACCTCCAGATTTCCAAAAAGAA	268
Qy	386	AACCTGCTAAAAGATACTCAAAAAGATGTGTGCACGCCAAAGATCTGACTGGCGACTAG	445
Db	269	AGCTCGAAAAAGAGACAGACGAGGAGGGAGGTGACAGAAAAGTGGAGAGACACTTGG	322
Qy	446	AGGTAGAGATAAAGATGCGATTTGAAAAATTTGAGCAAGAGAGACTGTAAAGCTCACAAAGC	505
Db	329	AAAAAGAGAGATAGAGGAGCAACGAAATATGCCCCAAAGAGCAACAGAGGTAAATGAAA	388
Qy	506	AACCAACAGCAAGATTGTAAACGACTATTAGACTTATGGGGGTTCTCTGTTTACAGGCAC	565
Db	389	TGCTCATCCAGAGATGCAAAAAACCTTAGAGCTTTTATGGGAAATTTCTTAATGTTTCAAGAC	448
Qy	566	CTTCTGACAGACAGACAGATGTGCAGCCCTTTGCATTAACAGATAAGAGTGTCCGTGTTG	625
Db	449	CTACCGAGGAGAGGCCCAAGTGTGATATTGCGCGCAAAAGGGAGACGTTATGTCATCGG	508
Qy	626	CTTCAGAGATTAAGCACTCCCTACTTTTGTGGGCTCCACGGTTCCTTGCTCATTTAATGG	683
Db	509	CTAGTCACAGATTTCGATTTCCCTACTTTTGTGGAGCTCCCAACACTGTATTGAACCTTAACAA	566
Qy	686	ATCCAACTTCCAGAAAATATCCGTGATGGAATTTGATTTG 727	
Db	569	TAAACAGAAAAAGAAAGTTGCTGTGGAAAAATTTGTCTACGTGG 610	

RESULT 12
 US-08-759-038-114
 : Sequence 114, Application US/08759038
 : Patent No. 6090543
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Prudent, James R.
 : APPLICANT: Hall, Jeff G.
 : APPLICANT: Lyamichev, Victor I.
 : APPLICANT: Brow, Mary Ann D.
 : APPLICANT: Dahlberg, James E.
 : TITLE OF INVENTION: Cleavage of Nucleic Acids
 : NUMBER OF SEQUENCES: 134
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Medien & Carroll, LLP
 : STREET: 220 Montgomery Street, Suite 2200
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States of America
 : ZIP: 94104
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/759,038
 : FILING DATE: 02-DEC-1996
 :
 : CLASSIFICATION:
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/
 : FILING DATE: 29-NOV-1996
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/682,853
 : FILING DATE: 12-JUL-1996
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/599,491
 : FILING DATE: 24-JAN-1996
 :
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ingolia, Diane E.
 : REGISTRATION NUMBER: 40,027
 : REFERENCE/DOCKET NUMBER: FOS-02574
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (415) 705-8410
 : TELEFAX: (415) 397-8338
 :
 : INFORMATION FOR SEQ ID NO: 114:

SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1020
US-08-759-038-114

Query Match 11.0%; Score 162.8; DB 3; Length 1023;
Best Local Similarity 55.0%; Pred. No. 2.1e-35;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 146 CGATGAGAGACGAGAGTTCGAGAGCTACTTGGCGCGCAAAATCGCGCTGACGCCAGCA 205
DB 29 CAAAGAAAGAAATTTGAGATTGAAAACCTATACGGAAAAAATCGCATCGACGCTCTTA 88
QY 206 TGACATCTCCAGTCTCTGATAGTAGTGGAGAGACGATGSAATCTCACAATG 265
DB 89 ATGCAATCTACCAATTTTGTTCACATTAAGACAGAAAGTGGAACTCCACTTAATGAT 148
QY 266 AAGCTGTGAAGTCACTAGTATTGGCAAGAAATGTTCAACCGACATATAGATTACTGG 325
DB 149 CAAAGGTAGATTAACCTCCACTAAGCGGCTCTTTTACAGACATTAACCTAATG 208
QY 326 AAGCGGAAATCAAGCGCTTATGTTTGTGATGGCAAGCTCCTGATATGAGAAAGCAAG 385
DB 209 AGGCTGGAATTAACCTGTATATGTTTGTATGATGAGAACTCCGAAATTCAAAGAAAG 268
QY 386 AACTTGCTAAAGATTAAGTCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGCAGTAG 445
DB 269 AGCTCGAAAAAAG 328
QY 446 AGTAGGAGATTAAGATGAGCATTTGAAAAATTTGAGCAAGAGCATGTAAGAGTCACAGGC 505
DB 329 AAAAAAGAGATTAAG 388
QY 506 AACCAACAGAGATTTGAAACGACTATTATAGACTTAATGGGGTTCCTGTTTGAAGGCAC 565
DB 389 TGTCTATCGAGAGATGCAAAAAAATCTTAGAGCTTAATGGAAATCTTAATAGTTCAAGC 448
QY 566 CTCTGAGAGAGAGAGAGATGAGAGCCCTTTGCAATTAACATTAAGTGTGCTGCTG 625
DB 449 CTACCGAGGAG 508
QY 626 CTTCAGAGATTAAGAGACTCCCTTACTTTTGGGCTCCAGCGTTCCTGCTCAATTTATG 685
DB 509 CTAGTCAAGATTAAGATTTCTTACTTTTGGAGCTCAAGACTGTGAATAACTTAACA 568
QY 686 ATCCAAGTCCAGAAATATACCTGTGATGGAATTTGATGCTG 727
DB 569 TAACAGAGAAAAAGAAAGTGTGCGGAAAAAATGTCTACGTCG 610

RESULT 13

US-08-758-314-114
Sequence 114, Application US/08758314
Patent No. 6090606

GENERAL INFORMATION:

APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Improved Cleavage Agents
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,314
FILING DATE: 02-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02575
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO.: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1020
US-08-758-314-114

Query Match 11.0%; Score 162.8; DB 3; Length 1023;
Best Local Similarity 55.0%; Pred. No. 2.1e-35;
Matches 320; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 146 CGATGAGAGAGAGAGATTCGAGAGCTACTTGGCGCGCAAAATCGCGCTGACGCCAGCA 205
DB 29 CAAAGAAAGAAATTTGAGATTGAAAACCTATACGGAAAAAATCGCATCGACGCTCTTA 88
QY 206 TGACATCTCCAGTCTCTGATAGTAGTGGAGAGACGAGATGSAATCTCACAATG 265
DB 89 ATGCAATCTACCAATTTTGTTCACATTAAGACAGAAAGATGGAATCTCAATTAATGAT 148
QY 266 AAGCTGTGAAGTCACTAGTATTGGCAAGAAATGTTCAACCGACATATAGATTACTGG 325
DB 149 CAAAGGTAGATTAACCTCCACTAAGCGGCTCTTTTACAGACATTAACCTAATG 208
QY 326 AAGCGGAAATCAAGCGCTTATGTTTGTGATGGCAAGCTCCTGATATGAGAAAGCAAG 385
DB 209 AGGCTGGAATTAACCTGTATATGTTTGTATGATGAGAACTCCAGAAATTCAGAAAGAG 268
QY 386 AACTTGCTAAAGATTAAGTCTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGCAGTAG 445
DB 269 AGCTCGAAAAAAG 328
QY 446 AGTAGGAGATTAAGATGAGCATTTGAAAAATTTGAGCAAGAGACTGTAAAGTCAACAGGC 505
DB 329 AAAAAAGAGATTAAG 388
QY 506 AACCAACAGAGATTTGAAACGACTATTATAGACTTAATGGGGTTCCTGTTTGAAGGCAC 565
DB 389 TGTCTATCGAGAGATGCAAAAAAATCTTAGAGCTTAATGGAAATCTTAATAGTTCAAGC 448
QY 566 CTCTGAGAGAGAGAGAGATGAGAGCCCTTTGCAATTAACATTAAGTGTGCTGCTG 625
DB 449 CTACCGAGGAG 508

Tue May 27 16:02:32 2003

us-09-805-311-7.std.rni

Page 13

Oy	626	CTTCGAGGATTAAGACACCCCTTACTTCTTGGGGGCTCCACAGCTTCCCTGCACTTTAATG	685
Db	509	CGAGTCAGATTATGATTCACACTCTTGTGTGCTCCAGTTCATTTAGCAATCTGCAA	568
Oy	686	ATCCAGTTCCTCCAGAAATACCTGTGATGCAGATTGATGTGG	727
Db	569	TTACGGGAAAAAGAAAGATGCTGGGAAGAGTGTTTACGTGG	610

```
Search completed: November 5, 2002, 11:03:14
Job time : 46.5823 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: November 5, 2002, 10:28:52 ; Search time 1264.61 Seconds
(without alignments)
15774.429 Million cell updates/sec

Title: US-09-805-311-7

Perfect score: 1478
Sequence: 1 cgaccacgcgtccgggaaa.....aaaaaaaaaaaaaaaa 1478

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 13736207 seqs, 674847542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_estrc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	793	53.7	901	10	BG837708 zml0_01f0
2	570	38.6	586	9	AI881599 606068G09
3	538.2	36.4	553	9	AM562789 660065H06
4	534.2	36.1	554	10	BE639421
5	530	35.9	550	10	BE639422
6	497.8	33.7	532	9	AM000375
7	471.4	31.9	474	9	AM559173
8	464	31.4	648	9	AV913663
9	454.4	30.7	467	9	BE186786
10	454	30.7	470	9	AI861468
11	451.8	30.6	475	9	AI861468
12	435	29.4	553	9	AM562517
13	417.2	28.2	733	9	AI834484
14	408.4	27.6	539	10	AM774700
15	399	27.0	414	9	BG263233
16	393.2	26.6	456	9	AM288784
17	375.6	25.4	632	9	AI065689
					AM585913

18	370.4	25.1	754	10	BE416584	BE416584	MUG009.B1
19	369.6	25.0	820	10	BG414505	BG414505	HVSMK000
20	369.4	25.0	609	10	BG096545	BG096545	EST461064
21	361.4	24.5	645	9	AM691681	AM691681	NE0428095
22	359.8	24.3	363	9	AM562518	AM562518	660065H06
23	348	23.5	680	10	BE417817	BE417817	MUG024.H1
24	347.2	23.5	683	9	AV918966	AV918966	AV918966
25	345.2	23.4	454	9	AV915644	AV915644	AV915644
26	345.2	22.4	539	10	BE098320	BE098320	EST428761
27	336	21.4	592	9	AT780966	AT780966	EST261845
28	314	21.2	470	9	AM288831	AM288831	707010F11
29	311.6	21.1	579	9	AV833393	AV833393	AV833393
30	310	21.0	376	9	AT065546	AT065546	a988e02.x
31	306.4	20.7	526	10	B1419720	B1419720	LJNEST479
32	295.8	20.0	695	9	BE052579	BE052579	CA_Pa000
33	291.4	18.4	549	9	AM696665	AM696665	NP109F01S
34	267.6	18.1	942	9	AL560007	AL560007	AL560007
35	259.4	17.6	903	9	AL531350	AL531350	AL531350
36	257.6	17.4	851	9	AL519300	AL519300	AL519300
37	243.8	16.5	689	10	BE311755	BE311755	601143566
38	240.6	16.3	769	10	BG574950	BG574950	602598038
39	238	16.1	660	10	BJ011893	BJ011893	BJ011893
40	238	16.1	806	9	AL560395	AL560395	AL560395
41	232.8	15.8	762	10	BG337603	BG337603	602435150
42	232	15.7	704	9	AM783795	AM783795	rp09g07.Y
43	229.8	15.5	866	10	BI827898	BI827898	603073821
44	229	15.5	812	10	BM015629	BM015629	603641864
45	228.6	15.5	963	12	CNS06WXR	AL419029	T3 end of

ALIGNMENTS

RESULT 1
LOCUS BG837708 901 bp mRNA linear EST 25-MAY-2001
DEFINITION zml0_01f08_A zml0_AAFRC_ECORC_Fusarium_graminearum_corn_silk zea
mays cDNA clone zml0_01f08, mRNA sequence.
ACCESSION BG837708.1 GI:14204031
VERSION BG837708
KEYWORDS EST.
SOURCE zea mays.
ORGANISM zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 901)
Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De Moors,R., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A., Sprott,D. and Tinker,N.A.
Expressed Sequence Tags from Maize Silk Six Hours After Silk Channel Inoculation with Fusarium graminearum

JOURNAL Unpublished (2001)
COMMENT Contact: Harris, Linda J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1314
Fax: (613) 759-6566
Email: harris@em.agr.ca.

FEATURES

source
1..901
/organism="Zea mays"
/cultivar="CO388"
/db_xref="taxon:4577"
/clone="zml0_01f08"
/clone_lib="zml0_AAFRC_ECORC_Fusarium_graminearum_corn_silk"
/tissue_type="Silk"
/dev_stage="4-5 days post-silk emergence"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI; Site_2: XhoI; Field-grown corn was silk channel-inoculated

in the morning (-10 am) with 1 ml of a macroconidial suspension (500,000 spores/ml) of *Fusarium graminearum* and silk channels were collected and immediately frozen in liquid nitrogen 6 hours later. RNA was extracted from silk tissue between 1 cm below and above the inoculation point in the silk channel, RNA from five silk channels was pooled."

BASE COUNT 267 a 174 c 208 g 249 t 3 others

ORIGIN

Query Match 53.7%; Score 793; DB 10; Length 901;
Best Local Similarity 98.6%; Pred. No. 3.5e-124;
Matches 82; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

QY 617 TCGCTGTTGCTTGAAGATAGAGTCCCTTACTTTGGGGCTCCAGGTTCCCTGTC 676
|||||
Db 1 TCGCTGTTGCTTGAAGATAGAGTCCCTTACTTTGGGGCTCCAGGTTCCCTGTC 60

QY 677 ATTAAATGATCCAGTTCACAAAATACCTGTGATGATTTGATGTTGCCAGGTTT 736
|||||
Db 61 ATTAAATGATCCAGTTCACAAAATACCTGTGATGATTTGATGTTGCCAGGTTT 120

QY 737 TGGAGAGCTTGAACCTCCATGACACAGTTCATTTGATCCTGTGTGATGTG 796
|||||
Db 121 TGGAGAGCTTGAACCTCCATGACACAGTTCATTTGATCCTGTGTGATGTG 180

QY 797 ACTATTTGATAGATCAAAAGTATGGGGGCAACACCTCGAATCTTATGCTCAC 856
|||||
Db 181 ACTATTTGATAGATCAAAAGTATGGGGGCAACACCTCGAATCTTATGCTCAC 240

QY 857 ATGGGTCATAGAAAGACTTGGAGATCTTAAATAAGACAGATATCAATTC -TGAG 915
|||||
Db 241 ATGGGTCATAGAAAGACTTGGAGATCTTAAATAAGACAGATATCAATTC -TGAG 300

QY 916 GACGAGCTTACC -AAGAACTGAGCTTGTTCAGAGAGCCATATGCTTGGATAT 974
|||||
Db 301 GACGAGCTTACC -AAGAACTGAGCTTGTTCAGAGAGCCATATGCTTGGATAT 360

QY 975 TCCAGCTTAAATGAGCTGACCTGATGAGAGGCTCTCAATGTTCTGTTAAAGA 1034
|||||
Db 361 TCCAGCTTAAATGAGCTGACCTGATGAGAGGCTCTCAATGTTCTGTTAAAGA 420

QY 1035 TAAATGTTCAATGAGATGGGGTCAAGGCTAGAGAGATCAATTCGCCAAGA 1094
|||||
Db 421 TAAATGTTCAATGAGATGGGGTCAAGGCTAGAGAGATCAATTCGCCAAGA 480

QY 1095 TAAATGTTCAAGGAGAGCTGAGTCTTTTCAAGCCACCTGCCACATCAGAC 1154
|||||
Db 481 TAAATGTTCAAGGAGAGCTGAGTCTTTTCAAGCCACCTGCCACATCAGAC 540

QY 1155 GCTAAAGGAGAGAGCTCGGATTAANCAAGCAGGAGCTGGACAAACAAACAA 1214
|||||
Db 541 GCTAAAGGAGAGAGCTCGGATTAANCAAGCAGGAGCTGGACAAACAAACAA 600

QY 1215 GCGTGTGGAAGAAGAAATATCTTGAGTCTTGATGATCAACTGAGACTGAG 1274
|||||
Db 601 GCGTGTGGAAGAAGAAATATCTTGAGTCTTGATGATCAACTGAGACTGAG 660

QY 1275 GCGGTGCAATGATCTCCCTAGATATTTTACCTCCTGTTTACCTCAACCTTTGG 1334
|||||
Db 661 GCGGTGCAATGATCTCCCTAGATATTTTACCTCCTGTTTACCTCAACCTTTGG 720

QY 1335 TGAAGTTTCCCATGTTTAAAGCTGGGTAAATGATGTTGTTGAAGATTTGGTGA 1394
|||||
Db 721 TGAAGTTTCCCATGTTTAAAGCTGGGTAAATGATGTTGTTGAAGATTTGGTGA 780

QY 1395 CCAAGTAAACAAATTTTCCCTG -TTTTTACTTCTTGCTTTCAGATATGATCCAG 1453
|||||
Db 781 CCAAGTAAACAAATTTTCCCTG -TTTTTACTTCTTGCTTTCAGATATGATCCAG 840

QY 1454 T 1454
+
Db 841 T 841

RESULT 2
LOCUS A1881599 586 bp mRNA linear EST 02-FEB-2000
DEFINITION 6060686G09.y1 606 - Ear tissue cDNA library from Schmidt lab Zea
mays cDNA, mRNA sequence.
ACCESSION A1881599
VERSION A1881599.1 GI:5566733
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 586)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1995)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 606068 row: 6 column: 09.
Location/Qualifiers
FEATURES
source
1..586
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="606 - Ear tissue cDNA library from Schmidt lab"
/tissue_type="mixed"
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOLR (Stratagene)"
/note="Organ: Immature ear; Vector: pBR-CMV; Site_1: EcoRI
; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt lab"

BASE COUNT 186 a 132 c 159 g 109 t

ORIGIN

Query Match 38.6%; Score 570; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GGAATAGCTCGGGGCGGGGTTTCTTGCCGACATCGGCTCAGCGCGCCGACACC 75
|||||
Db 17 GGAATAGCTCGGGGCGGGGTTTCTTGCCGACATCGGCTCAGCGCGCCGACACC 76

QY 76 GCCACAGCCGCCGACAGACAGATGGGCTTACGAAACTGCTGGCGACAT 135
|||||
Db 77 GCCACAGCCGCCGACAGACAGATGGGCTTACGAAACTGCTGGCGACAT 136

QY 136 GCCGCCAAGCGGTGAAGAGAGCAAGTTCGAGAGCTACTTGGCGCAAACTCTGCGGACAT 136
|||||
Db 137 GCCGCCAAGCGGTGAAGAGAGCAAGTTCGAGAGCTACTTGGCGCAAACTCTGCGGACAT 196

QY 196 GAGCGCAGCATGAGATCTACAGTTCTGATAGTATGATGGAAGACAGCATGAAATC 255
|||||
Db 197 GAGCGCAGCATGAGATCTACAGTTCTGATAGTATGATGGAAGACAGCATGAAATC 256

QY 256 CTCACAATGAACCTGCTGAAGCACTACTGATTTGCAAGGAAATGTTCAACGGGACATA 315
|||||
Db 257 CTCACAATGAACCTGCTGAAGCACTACTGATTTGCAAGGAAATGTTCAACGGGACATA 316

QY 316 AGATTACTGGAAGCGGAATCAAGCGATTATGTTTTGATGAGAGCTCTGTGATATG 375
|||||
Db 317 AGATTACTGGAAGCGGAATCAAGCGATTATGTTTTGATGAGAGCTCTGTGATATG 376

QY 376 AAGAACAAGAACTTCTTAAAGATATCTCAAAAAGATGATGCAACGAAGATCTGACT 435

```

Db 377 AAGAAACAGACTTGTCTAAAGATACCTCAAAAGAGATGATGACCAAGATCTACT 436
Oy 436 GAGCAGTAGAGTAGAGATGAATGCGATTTGAAAAATGAGCAGAGAGACTGTAAG 495
Db 437 GAGCAGTAGAGTAGAGATGAATGCGATTTGAAAAATGAGCAGAGAGACTGTAAG 496
Oy 496 GTCAACAGGCAACACAAAGAGATTGTAACGACTATTAGACTTTATGGGGCTTCGTT 555
Db 497 GTCAACAGGCAACACAAAGAGATTGTAACGACTATTAGACTTTATGGGGCTTCGTT 556
Oy 556 GTAGAGGCACTTCTGAAGCAGAGACGAA 585
Db 557 GTAGAGGCACTTCTGAAGCAGAGACGAA 586

RESULT 3
AW562789 553 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW562789
VERSION AW562789.1 GI:7216667
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade: Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 553)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNPUBLISHED (1999)
JOURNAL Contact: Walbot V
COMMENT Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
FEATURES
source
1..553
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 171 a 111 c 129 g 142 t
ORIGIN
Query Match 36.4%; Score 538.2; DB 9; Length 553;
Best Local Similarity 99.4%; Pred. No. 3.3e-81;
Matches 540; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 915 GCATGCGCTTACCAAGAGCTCGACGCTGTGTCAAGAGCGCTTAATGTCACTTTGGATAT 974
Db 4 GCATGCGCTTACCAAGAGCTCGACGCTGTGTCAAGAGCGCTTAATGTCACTTTGGATAT 63
Oy 975 TCCTAGCTTAATAGATGACCTGACCTGATGAGAGGGTCTCATAGTTCTCTGTGTAAGA 1034
Db 64 TCCTAGCTTAATAGATGACCTGACCTGATGAGAGGGTCTCATAGTTCTCTGTGTAAGA 123
Oy 1035 TATGCTTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAGAA 1094
Db 124 TATGCTTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAGAA 183

```

```

Oy 1095 TAAATGTCGCAAGAGAACTGAGTCCTTTTCAAGCCAACTGCCACCATACAGACC 1154
Db 184 TAAATGTCGCAAGAGAACTGAGTCCTTTTCAAGCCAACTGCCACCATACAGACC 243
Oy 1155 GCTAAAAAGGAGAACTTGGATTAACACAGCAGGACGCTGCCAACAAGAAAAA 1214
Db 244 GCTAAAAAGGAGAACTTGGATTAACACAGCAGGACGCTGCCAACAAGAAAAA 303
Oy 1215 GCGTGTGTAAGAAAGAAATATCTTGATGCTGTGATGATCAACATACGATAGAA 1274
Db 304 GCGTGTGTAAGAAAGAAATATCTTGATGCTGTGATGATCAACATACGATAGAA 363
Oy 1275 GCGGTGCATGATCACTTCCCTAGATTATTATCTACCTCCGTTTAACTCAGACCTTTGG 1334
Db 364 GCGGTGCATGATCACTTCCCTAGATTATTATCTACCTCCGTTTAACTCAGACCTTTGG 423
Oy 1335 TGAAGATTTGCCCATGTTTCAAGCTGGGTAAGTATGTTGTGTAAGAGATTGGTGA 1394
Db 424 TGAAGATTTGCCCATGTTTCAAGCTGGGTAAGTATGTTGTGTAAGAGATTGGTGA 483
Oy 1395 CCAAGTAAACAAACTATCGCTGTTTCTTCTGCTTGAAGTATGATGCGCAGT 1454
Db 484 CCAAGTAAACAAACTATCGCTGTTTCTTCTGCTTGAAGTATGATGCGCAGT 543
Oy 1455 AAA 1457
Db 544 ACA 546

RESULT 4
BE639421/c 554 bp mRNA linear EST 30-AUG-2000
LOCUS 946033A02.x2 946 - tassels primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION BE639421
VERSION BE639421.1 GI:9952838
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade: Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 554)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNPUBLISHED (1999)
JOURNAL Contact: Walbot V
COMMENT Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.
FEATURES
source
1..554
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassel primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybridZAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridZAP. Sample insert size range was 350 bp
to 3 kb with a 1 kb average."
BASE COUNT 145 a 131 c 110 g 168 t

```


ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 532)
 AUTHORS Walbot,V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 614014 row: D column: 03.

FEATURES
 source location/Qualifiers
 1..532
 /organism="Zea mays"
 /cultivar="W23"
 /db_xref="taxon:4577"
 /clone_lib="614 - root cDNA library from Walbot lab"
 /tissue_type="root"
 /dev_stage="3-4 days old"
 /lab_host="X10LR"
 /note="Organ: root; Vector: pBluescriptII SK+; Site_1:
 EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
 Lab (LM)"
 BASE COUNT 173 a 103 c 129 g 127 t

Query Match 33.7%; Score 497.8; DB 9; Length 532;
 Best Local Similarity 98.3%; Pred. No. 2.2e-74;
 Matches 524; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

OY 678 TTTATGATCCAAAGTTCCAGAAATACCTGATGATGATTTGATGTTGCCAAGGTTT 737
 Db 1 TTTATGATCCAAAGTTCCAGAAATACCTGATGATGATTTGATGTTGCCAAGGTTT 60
 OY 738 GGAGAGCTTGAAC 797
 Db 61 GGAGAGCTTGAAC 120
 OY 798 CTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
 Db 121 CTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 OY 858 TGGGTCCATAGAAAGCATCTTGGAGATCTTAATAAGACA-GATATCAAAATTCCTGAGG 916
 Db 181 TGGGTCCATAGAAAGCATCTTGGAGATCTTAATAAGAGATATCAAAATTCCTGAGG 240
 OY 917 ACTGGCCTTACCAAGAGCTGACGCTGTTCAGAGAGCTTATGCACTTTGGATATTC 976
 Db 241 ACTGGCCTTACCAAGAGCTGACGCTGTTCAGAGAGCTTATGCACTTTGGATATTC 299
 OY 977 CTGAGCTAAATGAGTGCACCTGATGAGAGGCTCTCAATGATTTCTCGTAAAGATA 1036
 Db 300 CTGAGCTAAATGAGTGCACCTGATGAGAGGCTCTCAATGATTTCTCGTAAAGATA 359
 OY 1037 ATGTTTCAATGAGATGCGGTGACAAAGGCAATAGAAAGCAATGCGCAAGATA 1096
 Db 360 ATGTTTCAATGAGATGCGGTGACAAAGGCAATAGAAAGCAATGCGCAAGATA 419
 OY 1097 AATCGTCGAAGAACTCGAGTCTTTTTCAGGCCAATGCGCACACATGAGACCGC 1156
 Db 420 AATCGTCGAAGAACTCGAGTCTTTTTCAGGCCAATGCGCACACATGAGACCGC 479
 OY 1157 TAAACGGAAGAGACTTGGATTAACCAAGCAAGGAGCTGCGAACAAGATA 1209
 Db 480 TAAACGGAAGAGACTTGGATTAACCAAGCAAGGAGCTGCGAACAAGATA 532

RESULT 7
 AM559173/c 474 bp mRNA linear EST 07-MAR-2000
 LOCUS 660065H06.xl 660 - Mixed stages of anther and pollen Zea mays cDNA,
 DEFINITION mRNA sequence.
 ACCESSION AM559173
 VERSION AM559173.1 GI:7204640
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 474)
 AUTHORS Walbot,V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660065 row: H column: 06.

FEATURES
 source location/Qualifiers
 1..474
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="X10LR"
 /note="Organ: anthers; Vector: Lambda Zap. Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."

BASE COUNT 113 a 114 c 92 g 154 t 1 others

Query Match 31.9%; Score 471.4; DB 9; Length 474;
 Best Local Similarity 99.6%; Pred. No. 6.2e-70;
 Matches 472; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 932 AAGCTGACGCTTGTTCAGAGAGCCTTAATGTCACCTTGGATATTCGAGCTAAATGGA 991
 Db 474 AAGCTGACGCTTGTTCAGAGAGCCTTAATGTCACCTTGGATATTCGAGCTAAATGGA 415
 OY 992 CTGCACCTGATGAGAGGCTCTCATAGATTCCTGTTAAAGATTAAGGTTTCAATGAAG 1051
 Db 414 CTGCACCTGATGAGAGGCTCTCATAGATTCCTGTTAAAGATTAAGGTTTCAATGAAG 355
 OY 1052 ATCGGGTGACAAAGCCATAGAGAAATCAAAATCTGCCAAGATTAATCGTCGAAGAA 1111
 Db 354 ATCGGGTGACAAAGCCATAGAGAAATCAAAATCTGCCAAGATTAATCGTCGAAGAA 295
 OY 1112 GACTCGAGTCTTTTTCAGAGCCACACGACCATGCGACCGCTAAAGAGGAAGAGA 1171
 Db 294 GACTCGAGTCTTTTTCAGAGCCACACGACCATGCGACCGCTAAAGAGGAAGAGA 235
 OY 1172 CTTCGGATTAACCAAGCAAGCAGCTGCGAACAAGAAACAAGGCTGGTGAAGAGA 1231
 Db 234 CTTCGGATTAACCAAGCAAGCAGCTGCGAACAAGAAACAAGGCTGGTGAAGAGA 175
 OY 1232 AATAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
 Db 174 AATAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 115
 OY 1292 TCGCCTAGATTAATTAACCTCTGTTTAACACACACTTTGGTGAAGATTTGCCATGT 1351
 Db 114 TCGCCTAGATTAATTAACCTCTGTTTAACACACACTTTGGTGAAGATTTGCCATGT 55

```

QY 1352 TTCACGCTGGGTAAGTCTGTTTGAAGAGATGCTGACCAAGTAAACA 1405
Db 54 TTCAAGCTGGGTAAGTCTGTTTGAAGAGATGCTGACCAAGTAAACA 1

RESULT 8
AV913663
LOCUS AV913663 648 bp mRNA linear EST 18-JAN-2002
DEFINITION AV913663 K. Sato unpublished cDNA library, cv. Haruna Niho
          germination shoots Hordeum vulgare subsp. vulgare cDNA clone
          bags22n17 5', mRNA sequence.
ACCESSION AV913663
VERSION AV913663.1 GI:18209440
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare.
          Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
          ; Triticeae; Hordeum.
          1 (bases 1 to 648)
REFERENCE Sato, K., Salsino, D. and Takeda, K.
AUTHORS Barley EST sequencing project in NIG and Okayama Univ
TITLE Unpublished (2002)
JOURNAL Contact: Tadasu Shin-i
          Center for Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshin@genes.nig.ac.jp.
FEATURES
  source
    1..648
    /organism="Hordeum vulgare subsp. vulgare"
    /cultivar="Haruna Niho"
    /db_xref="taxon:112509"
    /clone_lib="bags22n17"
    /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
    Niho germination shoots"
    /issue_type="shoots"
    /dev_stage="germination"

BASE COUNT 174 a 158 c 185 g 131 t
ORIGIN
Query Match 31.4%; Score 464; DB 9; Length 648;
Best Local Similarity 82.3%; Pred. No. 1e-68;
Matches 533; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

```

Db 361 AAGAGAGAGACAGCTTTTAAAAAGACGCAAGAGATGAGCAAGAGACTGACG 420
QY 436 GAGCAGTAGAGTAGAGATTAAGATGCGATTGAAAAATTGAGCAAGAGACTGTAAG 495
Db 421 AAGCAGTAGAGAGAGAGATACGATGCAATGAAATTCACAGAGAGACTGTAAAG 480
QY 496 GTACAGAGCAACAGACAGATGTAAGACGATTAATTAAGAGAGAGAGAGAGAGAG 555
Db 481 GTACAGAGCAACAGACAGATGTAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 556 GTAGAGGACAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
Db 541 GTAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 616 TTGCGTGTGCTTTCAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
Db 601 TATGCTTTGATCAGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648

RESULT 9
BE186786/c
LOCUS BE186786 467 bp mRNA linear EST 22-JUN-2000
DEFINITION 946012C08.X1 946 - tassal primordium prepared by Schmidt lab Zea
          mays cDNA, mRNA sequence.
ACCESSION BE186786
VERSION BE186786.1 GI:8665970
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
          clade; Panicoidae; Andropogoneae; Zea.
          1 (bases 1 to 467)
REFERENCE Walbot, V.
AUTHORS Maize ESTs from various cDNA libraries sequenced at Stanford
TITLE University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
          Department of Biological Sciences
          Stanford University
          855 California Ave, Palo Alto, CA 94304, USA
          Tel: 650 723 2227
          Fax: 650 725 8221
          Email: walbot@stanford.edu
          Plate: 946012 row: C column: 08.
FEATURES
  source
    1..467
    /organism="Zea mays"
    /cultivar="OH43"
    /db_xref="taxon:4577"
    /clone_lib="946 - tassal primordium prepared by Schmidt
    lab"
    /issue_type="tassels"
    /dev_stage="just after the transition from vegetative to
    inflorescence development"
    /lab_host="XLOLR"
    /note="Organ: tassels; Vector: HybridAP; Site 1: EcoRI;
    Site 2: XhoI; George Chnuck dissected immature tassels
    between 1mm and 3mm. Sharon Stanfield prepared the cDNA
    library in HybridAP. Sample insert size range was 350 bp
    to 3 kb with a 1 kb average."

BASE COUNT 108 a 110 c 96 g 153 t
ORIGIN
Query Match 30.7%; Score 454.4; DB 9; Length 467;
Best Local Similarity 99.6%; Pred. No. 4.6e-67;
Matches 466; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 958 AATGTCATTGGATATTTCCAGACTAAATGAGCTGCACCTGATGAGAGGGTCTCATATA 1017
 |||||||
 DB 407 AATGTCATTGGATATTTCCAGACTAAATGAGCTGCACCTGATGAGAGGGTCTCATATA 348
 QY 1018 AGTTTCCTGTAAGTAATGATGTTTCATGAAGATCGGTGACAAAGGCCATAGAGAG 1077
 |||||||
 DB 347 AGTTTCCTGTAAGTAATGATGTTTCATGAAGATCGGTGACAAAGGCCATAGAGAG 288
 QY 1078 ATCAATCTGCAGAAATTAATGCTGCCAAGAAAGTCTGATGCTTTTTCAGCCAACT 1137
 |||||||
 DB 287 ATCAATCTGCAGAAATTAATGCTGCCAAGAAAGTCTGATGCTTTTTCAGCCAACT 228
 QY 1138 GCCACCATCAGACCGCTTAAACGGAAGAGACTTCGATTAACGAAGAGCAGCT 1197
 DB 227 GCCACCATCAGACCGCTTAAACGGAAGAGACTTCGATTAACGAAGAGCAGCT 168
 QY 1198 GGAACAAGAAACAAAGGCTGTGGAAGAAAGAAATCTGTGATGCTGTGATGACAA 1257
 DB 167 GGAACAAGAAACAAAGGCTGTGGAAGAAAGAAATCTGTGATGCTGTGATGACAA 108
 QY 1258 CTACGACTAGCAAGACGCGTGGCATGATCATTGCGCTAGATTAATTAACCTCCTGTT 1317
 |||||||
 DB 107 CTACGACTAGCAAGACGCGTGGCATGATCATTGCG-CTAGATTAATTAACCTCCTGTT 49
 QY 1318 TTAACCTAGACCTTGGTGAAGTTTGCCCATGTTTCACAGCTGGGGTA 1365
 |||||||
 DB 48 TTAACCTAGACCTTGGTGAAGTTTGCCCATGTTTCACAGCTGGGGTA 1

RESULT 10
 A1861468/c 470 bp mRNA linear EST 19-JUL-1999
 LOCUS 614014D03.x1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
 DEFINITION mRNA sequence.
 ACCESSION A1861468
 VERSION A1861468.1 GI:5525575
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 470)
 Walbot, V.
 Zea mays ESTs from various cDNA libraries sequenced at Stanford

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 614014 row: D column: 03.
 Location/Qualifiers

FEATURES

source

1..470
 /organism="Zea mays"
 /cultivar="W23"
 /db_xref="taxon:4577"
 /clone_lib="614 - root cDNA library from Walbot Lab"
 /tissue_type="root"
 /dev_stage="3-4 days old"
 /lab_host="XLOLR"
 /note="Organ: root; Vector: pBlueScriptII SK+; Site_1:
 EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
 Lab (LM)"
 BASE COUNT 112 a 113 c 93 g 152 t
 ORIGIN

Query Match 30.7%; Score 454; DB 9; Length 470;
 Best Local Similarity 97.9%; Pred. No. 5,3e-67;
 Matches 460; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 949 AAGGACCTTAATCTACTTGGATATCTCTGAGCTAAATGAGACTGCACCTGATGAGAG 1008
 |||||||
 DB 470 AAGGACCTTAATCTCCCAATGATATCTCTGAGCTAAATGAGACTGCCCGCATGAGAGAG 411
 QY 1009 GGTCTCATAGTTCCTGTGTAAGATTAATGTTTCAATGAAGATCGGTGACAAAGGCC 1068
 |||||||
 DB 410 GGTCTCATAGTTCCTGTGTAAGATTAATGTTTCAATGAAGATCGGTGACAAAGGCC 351
 QY 1069 ATGAGAGATCTCAATCTGCGCAAGAAATATGCTGCGCAAGAGACTGAGCTTTTTC 1128
 DB 350 ATGAGAGATCTCAATCTGCGCAAGAAATATGCTGCGCAAGAGACTGAGCTTTTTC 291
 QY 1129 AAGCAACTCCACCATCAGCAGCGCTAAACGGAAGAGACTCGGATTAACAAAGC 1188
 DB 290 AAGCAACTCCCGCCCATCAGCAGCGCTAAACGGAAGAGACTCGGATTAACAAAGC 231
 QY 1189 AAGCAGCTCGCAACAAAGAAAGGCTGTGGAAGAGAAATATCTTGTGATGCTT 1248
 DB 230 AAGCAGCTCGCAACAAAGAAAGGCTGTGGAAGAGAAATATCTTGTGATGCTT 171
 QY 1249 GATGTACACTAGCAGCTAGCAAGAGCAGCGTGGCATGATCATTGCGCTAGATTAATTA 1308
 DB 170 GATGTACACTAGCAGCTAGCAAGAGCAGCGTGGCATGATCATTGCGCTAGATTAATTA 111
 QY 1309 CTCCTCTTTTACCTCAGACCTTTGGTGAAGTTTGCCCATGTTTCAGAGCTGGGTAACT 1368
 DB 110 CTCCTCTTTTACCTCAGACCTTTGGTGAAGTTTGCCCATGTTTCAGAGCTGGGTAACT 51
 QY 1369 TAGTTGTGTTGAAGATTTGGTGTACCAAGTAACAAACTTATGCTGT 1418
 |||||||

RESULT 11
 A1862517/c 475 bp mRNA linear EST 10-MAR-2000
 LOCUS 660065H06.x2 660 - Mixed stages of anther and pollen Zea mays cDNA,
 DEFINITION mRNA sequence.
 ACCESSION A1862517
 VERSION A1862517.1 GI:7216395
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 475)
 Walbot, V.
 Zea mays ESTs from various cDNA libraries sequenced at Stanford

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660065 row: H column: 06.
 Location/Qualifiers

FEATURES

source

1..475
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="XLOLR"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."

BASE COUNT 124 a 110 c 94 g 147 t
 ORIGIN

Query Match 30.6%; Score 451.8; DB 9; Length 475;
 Best Local Similarity 99.4%; Pred. No. 1.2e-66;
 Matches 464; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Y 991 ACTGACCTGATGAGGAGGCTCATAGTTCCGTGTAAGATATGTTTCATGTA 1050
 |||||||
 Db 475 ACTGACCTGATGAGGAGGCTCATAGTTCCGTGTAAGATATGTTTCATGTA 416
 Y 1051 GATCGGGTGACCAAGCCATAGAGATCAATCTGCCAAGATATATCGTCGAAGA 1110
 |||||||
 Db 415 GATCGGGTGACCAAGCCATAGAGATCAATCTGCCAAGATATATCGTCGAAGA 356
 Y 1111 AGACGCGAGTCCTTTTCAAGCCACAGCCACACATACACCGCTTAAGAGAGAG 1170
 |||||||
 Db 355 AGACTCGAGTCCTTTTCAAGCCACAGCCACATCAACCGCTTAAGAGAGAG 296
 Y 1171 ACTTCGATTAACACAGCAGCGTCGACAGCAAGAAACAAAGGCTGTGGAAGAAG 1230
 |||||||
 Db 295 ACTTCGATTAACACAGCAGCGTCGACAGCAAGAAACAAAGGCTGTGGAAGAAG 236
 Y 1231 AATATATCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1290
 |||||||
 Db 235 AATATATCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 176
 Y 1291 TTCGCTGATTAATTAACCTCCGTTTAACTACAGACCTTTGTGAAGTTGCCCATG 1350
 |||||||
 Db 125 TTCGCTGATTAATTAACCTCCGTTTAACTACAGACCTTTGTGAAGTTGCCCATG 116
 Y 1351 TTTCAGCTGGGTAAGTATGTTGTTGAAGATGTTGTTGTTGTTGTTGTTGTTGTT 1410
 |||||||
 Db 115 TTTCAGCTGGGTAAGTATGTTGTTGAAGATGTTGTTGTTGTTGTTGTTGTTGTT 57
 Y 1411 ATCCGCTGTTTACTCTTCTGCTTGAAGTATGTTGTTGTTGTTGTTGTTGTTGTT 1457
 |||||||
 Db 56 ATCCGCTGTTTACTCTTCTGCTTGAAGTATGTTGTTGTTGTTGTTGTTGTTGTT 10

RESULT 12
 A1834484/c 553 bp mRNA linear EST 02-FEB-2000
 LOCUS 60606609.xl 606 - Ear tissue cDNA library from Schmidt lab Zea
 DEFINITION
 mays cDNA, mRNA sequence.
 ACCESSION A1834484
 VERSION A1834484.1 GI:5468693
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 553)
 AUTHORS Walbot,V.
 TITLE Mays ESTs from various cDNA libraries sequenced at Stanford
 University
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 606068 row: G column: 09.

FEATURES

Source
 1. 553
 /organism="Zea mays"
 /culturivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="606 - Ear tissue cDNA library from Schmidt
 lab"

/tissue_type="mixed"
 /dev_stage="ear length from 0.5 cm - 2.0 cm"
 /lab_host="XLOLR (Stratagene)"
 /note="Organ: Immature ear; Vector: pBK-CMV; Site_1: EcoRI
 ; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
 lab"

BASE COUNT 140 a 112 c 106 g 195 t
 ORIGIN

Query Match 29.4%; Score 435; DB 9; Length 553;
 Best Local Similarity 97.8%; Pred. No. 8.1e-64;
 Matches 441; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Y 1007 AGGTCATCAATGTTCCGTGTAAGATATGTTTCATGATGATGATGATGATGATGAT 1066
 |||||||
 Db 553 AGGTCATCAATGTTCCGTGTAAGATATGTTTCATGATGATGATGATGATGATGAT 494
 Y 1067 CCATAGGAAGATCAATCTGCCAAGATTAATCTGCGAAGAGACTCGAGTCCTTT 1126
 |||||||
 Db 493 CCATAGGAAGATCAATCTGCCAAGATTAATCTGCGAAGAGACTCGAGTCCTTT 434
 Y 1127 TCAAGCCAACTGCCACACATGACACCGCTAAACGGAAGAGACTTCGGATAAACAA 1186
 |||||||
 Db 433 TCAAGCCAACTGCCACACATGACACCGCTAAACGGAAGAGACTTCGGATAAACAA 374
 Y 1187 GCAAGGAGCTGGGACACAGAAACAAAGGCTGTGGAAGAAGAAATATCTTGATGC 1246
 |||||||
 Db 373 GCAAGGAGCTGGGACACAGAAACAAAGGCTGTGGAAGAAGAAATATCTTGATGC 314
 Y 1247 TTGATGTACACTACGATACGATACGAAAGCAGCGGTGATGATGATGATGATGAT 1306
 |||||||
 Db 313 TTGATGTACACTACGATACGATACGAAAGCAGCGGTGATGATGATGATGATGAT 254
 Y 1307 AACCTCCGTTTAACTGACAGCTTTGTTGAAGTTGGCCATGTTTCAAGTGGGGTAA 1366
 |||||||
 Db 253 AACCTCCGTTTAACTGACAGCTTTGTTGAAGTTGGCCATGTTTCAAGTGGGGTAA 194
 Y 1367 GTTAGTGTGTTTGAAGAGATGTTGTTTACCAAGTAAACAACCTATTCGCTTTTACT 1426
 |||||||
 Db 193 GTTAGTGTGTTTGAAGAGATGTTGTTTACCAAGTAAACAACCTATTCGCTTTTACT 134
 Y 1427 TCTTGCTCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1457
 |||||||
 Db 133 TCTTGCTCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 103

RESULT 13
 AM774700 733 bp mRNA linear EST 07-SEP-2000
 LOCUS EST333851 KV3 Medicago truncatula cDNA clone PKV3-23L21, mRNA
 DEFINITION
 sequence.
 ACCESSION AM774700
 VERSION AM774700.1 GI:7718617
 KEYWORDS EST.
 SOURCE Medicago truncatula
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 REFERENCE 1 (bases 1 to 733)
 AUTHORS Vandenbosch,K., Hurl,J., Moore,J., Beremand,P., Peng,H., Ellis,L.,
 Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,J.E. and
 Fraser,C.M.
 TITLE ESTs from roots of Medicago truncatula after Rhizobium inoculation
 JOURNAL Unpublished (1999)
 COMMENT Contact: Vandenbosch K
 Department of Biology
 Texas A&M University
 College Station, TX 77843-3258, USA
 Tel: 409 845 7707
 Fax: 409 845 2891
 Email: Kate@mail.bio.tamu.edu

Db	121	CABA	CAATGTTTCACCGGACATATAGTGTCTCGACGAGGAGAAATTAAACAGTATATGTT	180
QY	352	TTTGATGGCAAGCCTCCTGATATGAGAAACACAGACTTGTCTAAAGATFACTCAAAAAGA	411	
Db	181	TTTATGGCAAGCCTCCTGAAATGAGAAAGGACGCGCTTTTAAAGACAGCAAGAGG	240	
QY	412	GATGATGCAACCAAGATCTGACTGAGCAGTAGAGTAGAGATTAAGATGCGATTGAA	471	
Db	241	AATGAGCAGACAGAGAGCTGACCAAGGACAGTAGAGAGATACGATATGCAATTGAA	300	
QY	472	AAATGAGCAGAGAGACTGTAAAGGTCACAAGCAACAACAGAGATTGTAAACGACTA	531	
Db	301	AAATTCACACAGAGAACTGTAAAGTCCAGGACGACGACCAATGATGATTGAAGCGCTTA	360	
QY	532	TTAAGACTTAAGGGGCTTCCTGCTCTGCTTAAGGACACTTCCTGAAAGCAGCAAGATGTGCA	591	
Db	361	CTTAAGCTTAAGGGGCTTCCTGCTCTGCTTAAGGACACTTCCTGAAAGCAGCAAGATGTGCT	420	
QY	592	GCCCTTTGCATAAACGATAGAGTGTTCCTGCTTCCTGCTGCAAGATAGAGACTCCCTTACT	651	
Db	421	GCCCTTTGCAGAAATGACAGAGGTGATCTGTTCATCGCAAGATATGAGTACTTACT	480	
QY	652	TTTGGGCTCCACGCTTCTTCCTGATTTATGATCCAGTCCAGAAATATACTCTGT	710	
Db	481	TTTGGAGCTACACGCTTGTGCTGCTATTTGATGATGATTCAGAGTCCAGAAATATACTCTGT	539	
RESULT 15				
AM288784	LOCUS	AM288784	414 bp	mRNA linear EST 16-JAN-2000
DEFINITION	707010C02.x5 707 - Mixed adult tissues from walbot lab (Sk) Zea			
ACCESSION	AM288784			
VERSION	AM288784.1	GI:6695706		
KEYWORDS	EST			
SOURCE	Zea mays.			
ORGANISM	Zea mays.			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC			
TITLE	clade; Panicoidae; Andropogoneae; Zea.			
JOURNAL	1 (bases 1 to 414)			
COMMENT	Walbot, V.			
	Maize ESTs from various cDNA libraries sequenced at Stanford			
	University			
	Unpublished (1999)			
	Contact: Walbot V			
	Department of Biological Sciences			
	Stanford University			
	855 California Ave, Palo Alto, CA 94304, USA			
	Tel: 650 723 2227			
	Fax: 650 725 8221			
	Email: walbot@stanford.edu			
	Plate: 707010 row: C column: 02.			
FEATURES	Location/Qualifiers			
source	1..414			
	/organism="Zea mays"			
	/cultivar="W23"			
	/db_xref="taxon:4577"			
	/clone_lib="707 - Mixed adult tissues from Walbot lab (Sk)"			
	/tissue_type="tassel, kernel, silk, husk, root, leaf"			
	/dev_stage="adult"			
	/lab_host="DH10B"			
	/note="Organ: tassel, kernel, silk, husk, root, leaf;			
	vector: pGAD10; Site1: EcorI; cDNA library from fully			
	differentiated maize tissues from an active Mutator			
	plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,			
	husk, root, leaf). Unidirectionally cloned."			
BASE COUNT	120 a 82 c 101 g 109 t 2 others			
ORIGIN				
Query Match	27.0%;	Score 399;	DB 9;	Length 414;
Best Local Similarity	98.3%;	Pred No. 1e-57;		

	Matches	402,	Conservative	0;	Mismatches	7,	Indels	0;	Gaps	0;
QY	661	CCACGGTCCCTTCGTCATTTAATGATGCCAAGTTCGCAAGAAATACCTGATGATGCAATTT	720							
Db	6	CCACGGTCCCTTCGTCATTTAATGATGCCAAGTTCGCAAGAAATACCTGATGATGCAATTT	65							
QY	721	GATGTTGCCAAGTTTGGAGAGGTTGACACACATGACGACCGACGTCATTGATTTGTGC	780							
Db	66	GATGTTGCCAAGTTTGGAGAGGTTGACACACATGACGACCGACGTCATTGATTTGTGC	125							
QY	781	ATCCGTGTGGATGTGACTATTGTGATAGCATCAAAGTATCGGGGGCCAAACAGCTCTG	840							
Db	126	ATCCGTGTGGATGTGACTATTGTGATAGCATCAAAGTATCGGGGGCCAAACAGCTCTG	185							
QY	841	AAACTTATTCGTAAACATGGGTCCATACGAAGCATCTTGGAGATCTTAATAAAGACAGA	900							
Db	186	AAACTTATTCGTAAACATGGGTCCATACGAAGCATCTTGGAGATCTTAATAAAGACAGA	245							
QY	901	TATCAATTCCTGAGACTGCGCTTACGACAGAAAGCTCGACCCCTTGTTCGAAGAGCCTAAT	960							
Db	246	TACCAATTCCTGAGACTGCGCTTACGACAGAAAGCTCGACCCCTTGTTCGAAGAGCCTAAT	305							
QY	961	GTCACTTTGGATATTCCTAGACTAAAATGAGACTGCACCTGATGAGGAGGCTTCATAAGT	1020							
Db	306	GTCACATTTGATGTTCCTGAGCTAAATATGACTGCACTGTATGAGGAGGCTTCATAAGT	365							
QY	1021	TTCCTGTAAAGATATATGTTTCAATGAGAAATGGGAGACAAAGGCCA	1069							
Db	366	TTCCTGTAAAGATATATGTTTCAATGAGAAATGGGAGACAAAGGCCA	4							

Search completed: November 5, 2002, 13:37:11
Job time : 1277.61 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 11:03:23 : Search time 1855.32 Seconds
(Without alignments)
16501.458 Million cell updates/sec

Title: US-09-805-311-1
Perfect score: 1463

Sequence: 1 cagcagatactcgcgcgc.....aaaaaaaaaaaaaaaa 1463

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapept 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank1 :
1: gb_ba :
2: gb_bt :
3: gb_in :
4: gb_cm :
5: gb_ov :
6: gb_pc :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vl :
15: em_da :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :
29: em_vl :
30: em_htg_hum :
31: em_htg_inv :
32: em_htg_other :
33: em_htgo_inv :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB	ID	Description
------------	-------------	--------	----	----	-------------

1	1463	100.0	1463	6	ARI52403	ARI52403 Sequence
2	1360	93.0	1541	6	ARI52404	ARI52404 Sequence
3	1267	86.6	1381	6	ARI52405	ARI52405 Sequence
4	894	61.1	1478	6	ARI52406	ARI52406 Sequence
5	58	4.0	1354	8	AB021666	AB021666 Oryza sat
6	58	4.0	115907	2	AC104713	AC104713 Oryza sat
7	43	2.9	105137	2	HS104017	HS104017 Human DNA
8	40	2.7	2871	9	BC022967	BC022967 Homo sapi
9	39	2.7	89181	2	AC084072	AC084072 Homo sapi
10	39	2.7	288	6	AX185400	AX185400 Sequence
11	39	2.7	305	6	AX186826	AX186826 Sequence
12	39	2.7	224448	2	PFMAL4P4	PFMAL4P4 Plasmid
13	38	2.6	659	6	BC009510	BC009510 Homo sapi
14	38	2.6	691	6	AX011676	AX011676 Sequence
15	38	2.6	2228	9	HS0802710	HS0802710 Homo sapi
16	38	2.6	2240	4	AB042653	AB042653 Bos tauri
17	38	2.6	2291	9	BC001537	BC001537 Homo sapi
18	38	2.6	3315	3	AY069109	AY069109 Drosophila
19	38	2.6	67632	2	AC105102	AC105102 Homo sapi
20	38	2.6	38734	2	PFMAL1P2	PFMAL1P2 Plasmid
21	38	2.6	119077	9	AL354962	AL354962 Human DNA
22	38	2.6	134963	9	HS199116	HS199116 Human DNA
23	38	2.6	145069	2	AC012032	AC012032 Homo sapi
24	38	2.6	165894	9	AC092054	AC092054 Homo sapi
25	38	2.6	196560	2	AC027646	AC027646 Mus muscu
26	38	2.6	203250	9	AC020910	AC020910 Homo sapi
27	38	2.6	238728	2	AC025528	AC025528 Mus muscu
28	37	2.5	396	6	AX093307	AX093307 Sequence
29	37	2.5	445	9	BC009480	BC009480 Homo sapi
30	37	2.5	508	6	E01762	E01762 Variant aeg
31	37	2.5	508	6	E01765	E01765 DNA sequenc
32	37	2.5	826	6	E08516	E08516 DNA encodin
33	37	2.5	826	6	E12586	E12586 cDNA encodi
34	37	2.5	826	6	S65048	S65048 self-incorp
35	37	2.5	835	8	AF064030	AF064030 Helianthu
36	37	2.5	835	8	ATP27A	ATP27A
37	37	2.5	848	10	BC018351	BC018351 Mus muscu
38	37	2.5	925	3	AEVA0440X	L29571 Aequorea vi
39	37	2.5	958	6	ARI08851	ARI08851 Sequence
40	37	2.5	958	6	ARI19320	ARI19320 Sequence
41	37	2.5	958	6	ARI151721	ARI151721 Sequence
42	37	2.5	958	6	AX250563	AX250563 Sequence
43	37	2.5	958	6	E01573	E01573 cDNA sequen
44	37	2.5	958	6	E01575	E01575 cDNA sequen
45	37	2.5	958	6	E02319	E02319 DNA encodin

ALIGNMENTS

RESULT 1
LOCUS ARI52403
DEFINITION Sequence 1 from patent US 6232527.
ACCESSION ARI52403
VERSION ARI52403.1 GI:15118453
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1463)
AUTHORS Manajan, P.B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 1 15-MAY-2001;
FEATURES
source location/Qualifiers
1..1463
BASE COUNT 466 a 292 c 361 g 344 t
ORIGIN

Query Match	Score	DB	Length
Best local Similarity	100.0%	100.0%	1463
Matches	1463	Conservative	0
Mismatches	0	Indels	0
Gaps	0	Gaps	0

QY 1 CAGAGATATAGCTCGGGCGCGGTTTCTTGCCCACTCCGGCTACGCCCGCCGCCCA 60
1 CAGAGATATAGCTCGGGCGCGGTTTCTTGCCCACTCCGGCTACGCCCGCCGCCCA 60
Db 1 CAGAGATATAGCTCGGGCGCGGTTTCTTGCCCACTCCGGCTACGCCCGCCGCCCA 60
QY 61 CCGGCCACAGCGCGCCGACATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 120
61 CCGGCCACAGCGCGCCGACATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 120
Db 61 CCGGCCACAGCGCGCCGACATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 120
QY 121 AATGCCCGCAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
121 AATGCCCGCAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 AATGCCCGCAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 GTGAGCGCAGCATGAGATATACAGTTCGATTTGATTTGATTTGATTTGATTTGAT 240
181 GTGAGCGCAGCATGAGATATACAGTTCGATTTGATTTGATTTGATTTGATTTGAT 240
Db 181 GTGAGCGCAGCATGAGATATACAGTTCGATTTGATTTGATTTGATTTGATTTGAT 240
QY 241 ACTCTACCAATGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 300
241 ACTCTACCAATGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 241 ACTCTACCAATGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 AATAAGATTAAGTGAAGCGGAGATCAAGCGAGTTTATTTTATTTTATTTTATTTTAT 360
301 AATAAGATTAAGTGAAGCGGAGATCAAGCGAGTTTATTTTATTTTATTTTATTTTAT 360
Db 301 AATAAGATTAAGTGAAGCGGAGATCAAGCGAGTTTATTTTATTTTATTTTATTTTAT 360
QY 361 ATGAGAGAGAGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 420
361 ATGAGAGAGAGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 361 ATGAGAGAGAGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 ACTGAGCGATAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 480
421 ACTGAGCGATAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 480
Db 421 ACTGAGCGATAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 480
QY 481 AAGGTACAG 540
481 AAGGTACAG 540
Db 481 AAGGTACAG 540
QY 541 GTTGTAG 600
541 GTTGTAG 600
Db 541 GTTGTAG 600
QY 601 GTTGTAG 660
601 GTTGTAG 660
Db 601 GTTGTAG 660
QY 661 CGTCAATTAATGAGATCAAGAGTTCAGAGAAATACCTGTGATGAGATGATGATGAT 720
661 CGTCAATTAATGAGATCAAGAGTTCAGAGAAATACCTGTGATGAGATGATGATGATGAT 720
Db 661 CGTCAATTAATGAGATCAAGAGTTCAGAGAAATACCTGTGATGAGATGATGATGATGAT 720
QY 721 GTTGTAG 780
721 GTTGTAG 780
Db 721 GTTGTAG 780
QY 781 TGTGACTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
781 TGTGACTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 TGTGACTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 CAGCATGAG 900
841 CAGCATGAG 900
Db 841 CAGCATGAG 900
QY 901 GAG 960
901 GAG 960
Db 901 GAG 960
QY 961 ATTCTGAG 1020
961 ATTCTGAG 1020
Db 961 ATTCTGAG 1020
QY 1021 GATTAATGTTTCAAG 1080
1021 GATTAATGTTTCAAG 1080
Db 1021 GATTAATGTTTCAAG 1080

QY 1081 AATAATGTCGCAAG 1140
1081 AATAATGTCGCAAG 1140
Db 1081 AATAATGTCGCAAG 1140
QY 1141 CCGCTAAACGGAAG 1200
1141 CCGCTAAACGGAAG 1200
Db 1141 CCGCTAAACGGAAG 1200
QY 1201 AAGGCTGTGGAAG 1260
1201 AAGGCTGTGGAAG 1260
Db 1201 AAGGCTGTGGAAG 1260
QY 1261 CAGGCTGTGGAAG 1320
1261 CAGGCTGTGGAAG 1320
Db 1261 CAGGCTGTGGAAG 1320
QY 1321 GGTAAAGTTTGTCTCATGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
1321 GGTAAAGTTTGTCTCATGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 GGTAAAGTTTGTCTCATGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 TACCAAGTAAACAAACTATGCTGTGTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
1381 TACCAAGTAAACAAACTATGCTGTGTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
Db 1381 TACCAAGTAAACAAACTATGCTGTGTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
QY 1441 AAAAAAAAAAAAAAAAAAAAAA 1463
1441 AAAAAAAAAAAAAAAAAAAAAA 1463
Db 1441 AAAAAAAAAAAAAAAAAAAAAA 1463
RESULT 2
AR152404 1541 bp DNA linear PAT 08-AUG-2001
LOCUS AR152404
DEFINITION Sequence 3 from patent US 6232527.
ACCESSION AR152404
VERSION AR152404.1 GI:15118454
KEYWORDS
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1541)
AUTHORS Mahajan, P. B.
TITLE Maize Red2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 3 15-MAY-2001;
FEATURES
source
1. AR1541
/organism="unknown"
BASE COUNT 473 a 308 c 377 g 383 t
ORIGIN
Query Match 93.0%; Score 1360; DB 6; Length 1541;
Best Local Similarity 99.9%; Pred. No 0;
Matches 1410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 GCGGTTTCTTGCGCCACTCGGCTCAGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGAG 80
21 GCGGTTTCTTGCGCCACTCGGCTCAGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGAG 80
Db 15 GCGGTTTCTTGCGCCACTCGGCTCAGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGAG 74
QY 81 CGAGATGGGATCAAGAGGTTTGAGCAAACTGCTGGCGGAGCAATGCGGCCAAGGCGATGAA 140
81 CGAGATGGGATCAAGAGGTTTGAGCAAACTGCTGGCGGAGCAATGCGGCCAAGGCGATGAA 140
Db 75 CGAGATGGGATCAAGAGGTTTGAGCAAACTGCTGGCGGAGCAATGCGGCCAAGGCGATGAA 134
QY 141 GGAGCAGAAATTCGAGAGTCTGCGCGCCAAATTCGCGCTGAGAGCGCAGCATGAGCAT 200
141 GGAGCAGAAATTCGAGAGTCTGCGCGCCAAATTCGCGCTGAGAGCGCAGCATGAGCAT 200
Db 135 GGAGCAGAAATTCGAGAGTCTGCGCGCCAAATTCGCGCTGAGAGCGCAGCATGAGCAT 194
QY 201 ATTCCATTCCTGATTTGATTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 260
201 ATTCCATTCCTGATTTGATTTGGAAG 260
Db 195 ATTCCATTCCTGATTTGATTTGGAAG 254
QY 261 TGAAGTCACTAGTCAATTTCAAG 320
261 TGAAGTCACTAGTCAATTTCAAG 320
Db 255 TGAAGTCACTAGTCAATTTCAAG 314

OY	321	AATCAAGCAGTTTATGTGTTTTGGTGAACACCCTCCGTGATATGAAAGAACAAAGACTTCG	380
Db	315	AATCAAGCAGTTTATGTGTTTTGGTGAACACCCTCCGTGATATGAAAGAACAAAGACTTCG	374
OY	381	TAAAAGATTACTCAAAAAAGAGATGATGCACAACCAAAGATCTACTAGAGCACTAAGGTAG	440
Db	375	TAAAAGATTACTCAAAAAAGAGATGATGCACAACCAAAGATCTACTAGAGCACTAAGGTAG	434
OY	441	AATATAAGATCGGATTTGAAAAATTGACGAGAAGGACTGTAAAGGTCCAAGCACACAA	500
Db	435	AATATAAGATCGGATTTGAAAAATTGACGAGAAGGACTGTAAAGGTCCAAGCACACAA	494
OY	501	CGAAGATTGTAAAGCGCTATTACACTTAATGGGGTTCCTGTTGPAGGACCTTGCA	560
Db	495	CGAAGATTGTAAAGCGCTATTAAAGCTTATGGGGTTCCTGTTGPAGGACCTTGCTGA	554
OY	561	AGCAGAGCAGATGTGCAGCCCTTTGCCATTAACGATMAAGGTGTGCTGTCTCAGA	620
Db	555	AGCAGAGCAGATGTGCAGCCCTTTGCCATTAACGATMAAGGTGTGCTGTCTCAGA	614
OY	621	AGATATGGACCTCCCTACTTTTGGGGGCTCCAGCGTTCCTCGCATTTAATGATCCAG	680
Db	615	AGATATGGACCTCCCTACTTTTGGGGGCTCCAGCGTTCCTCGCATTTAATGATCCAG	674
OY	681	TTTCCAAGAAAAATACCTGTGATGGAATTTGATGTGGCCAGAGTTTGGAGAGCTTGAACT	740
Db	675	TTTCCAAGAAAAATACCTGTGATGGAATTTGATGTGGCCAGAGTTTGGAGAGCTTGAACT	734
OY	741	CACCATGGACCAAGTTCATTGATTTGTGCACTCTGTGTGATGTGACTATTGTGATAGCAT	800
Db	735	CACCATGGACCAAGTTCATTGATTTGTGCACTCTGTGTGATGTGACTATTGTGATAGCAT	794
OY	801	CAAAAGTATCGGGGGGGCCAAACAGCTCGAAACCTATTCGACACATGGGGCCATAGAAAG	860
Db	795	CAAAAGTATCGGGGGGGCCAAACAGCTCGAAACCTATTCGACACATGGGTCCATAGAAAG	854
OY	861	CATCTTGGAGATCTTAATTAAGAACAGATATCAAAATTCCTGAGAGCTGGCTTACCAAGA	920
Db	855	CATCTTGGAGATCTTAATTAAGAACAGATATCAAAATTCCTGAGAGCTGGCTTACCAAGA	914
OY	921	AGCTGAGCGCTTGTTCAAGAGACCTTAATGACATTTGCATTTGCGAGCTTAAATGAGAC	980
Db	915	AGCTGAGCGCTTGTTCAAGAGAGCTTAATGACATTTGCATTTGCGAGCTTAAATGAGAC	974
OY	981	TGCACCTGATGAGAGAGGCTCATAAAGTTTCTGTGTAAGAGATMAATGTTTCAACGAGA	1040
Db	975	TGCACCTGATGAGAGAGGCTCATAAAGTTTCTGTGTAAGAGATMAATGTTTCAACGAGA	1034
OY	1041	TGCGGTGCAAAAGGCAATAGGAAGATCAAAATGTGCANAATAATTCGTCCCAAGAG	1100
Db	1035	TGCGGTGCAAAAGGCAATAGGAAGATCAAAATGTGCANAATAATTCGTCCCAAGAG	1094
OY	1101	ACTGAGTCTTTTTTCAAGCCAACTGCGACACATCAGACCGCTTAAACGAGAGAGAC	1160
Db	1095	ACTGAGTCTTTTTTCAAGCCAACTGCGACACATCAGACCGCTTAAACGAGAGAGAC	1154
OY	1161	TTTGGATTAAMCAAGCAAGGACGTGTGGAAACAAAGAAACAAAGGCTGTGTGAAAGAGAA	1220
Db	1155	TTTGGATTAAMCAAGCAAGGACGTGTGGAAACAAAGAAACAAAGGCTGTGTGAAAGAGAA	1214
OY	1221	ATAATCTTGATGCTGTGATGTATACAACTAGCACTACGAAAGCAAGCGGTGGCTGATCACTT	1280
Db	1215	ATAATCTTGATGCTGTGATGTATGACAACTAGCACTACGAAAGCAAGCGGTGGCTGATCACTT	1274
OY	1281	CGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTGGTAAAGTTTGCTCATGTT	1340
Db	1275	CGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTGGTAAAGTTTGCTCATGTT	1334
OY	1341	TCAAGCTGGGGTAGTGAATGTGTTTGAAGAGATGGGTGACCAAGTAAGCAAAACTTAT	1400
Db	1335	TCAAGCTGGGGTAGTGAATGTGTTTGAAGAGATGGGTGACCAAGTAAGCAAAACTTAT	1394
OY	1401	CGCTGTTTTTACTCTTGCTGTGCTTTGAAGTA	1431

Db	1395	CGCTGTTTTTACTTCTTGTCTTGAAGTA	1425
RESULT 3			
ARI52405	ARI52405	1381 bp	DNA
LOCUS	Sequence 5 from patent US 6232527.		Linear
DEFINITION			PAT 08-AUG-2001
ACCESSION	ARI52405		
VERSION	ARI52405.1	GI:15118455	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1381)		
AUTHORS	Mahajan, P. B.		
TITLE	Maize Rad2/FEN-1 orthologues and uses thereof		
JOURNAL	Patent: US 6232527-A 5 15-MAY-2001;		
FEATURES	Location/Qualifiers		
source	1..1381		
BASE COUNT	441 a 269 c 346 g 325 t		
ORIGIN			
Query Match	86.6%; Score 1267; DB 6; Length 1381;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 1267; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
64	GCACACGCCGCCGACAGACAGATGGGCATCAAGGCTTTGACGAACCTCTGGCGACAAAT	123	
Db	16	GCACACGCCGCCGACAGACAGATGGGCATCAAGGCTTTGACGAACCTCTGGCGACAAAT	75
64	GCGCCCAAGCGCATGAAGAGCAGAAAGTTCGAGAGCTCTTGCGCCGGAATAAGCGCGTC	183	
Db	76	GGGCCCAAGCGCATGAAGAGCAGAAAGTTCGAGAGCTCTTGCGCCGGAATAAGCGCGTC	135
184	GACGCCAGCATGAGCATATACCAAGTTCCTGATTTGAGTTGGAAGGACAGGCGTGAACACT	243	
Db	136	GACGCCAGCATGAGCATATACCAAGTTCCTGATTTGAGTTGGAAGGACAGGCGTGAACACT	195
244	CTCACAATGAAGCGTGGTGAAGTCACTAGTCTTTGCAAGGAATGTTCAACGGGCAATA	303	
Db	196	CTCACAATGAAGCGTGGTGAAGTCACTAGTCTTTGCAAGGAATGTTCAACGGGCAATA	255
304	AGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTTGATGGCAAGCCTCTGATATG	363	
Db	256	AGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTTGATGGCAAGCCTCTGATATG	315
364	AGGAACAAGACCTGCTAAAGATACGTCAAAAAGATGATGCACCAAGATCTGACT	423	
Db	316	AGGAACAAGACCTGCTAAAGATACGTCAAAAAGATGATGCACCAAGATCTGACT	375
424	GAGGACGTAGAAGTGAAGATTAAGATCCGATTGAAAATTGAGCAAGAGACGTAAAG	483	
Db	376	GAGGACGTAGAAGTGAAGATTAAGATCCGATTGAAAATTGAGCAAGAGACGTAAAG	435
484	GTCACAAGGCACAACAACGAAGATTGTAAAGCGCTATTAAAGCTTATGCGGCTCTGTTT	543	
Db	436	GTCACAAGGCACAACAACGAAGATTGTAAAGCGCTATTAAAGCTTATGCGGCTCTGTTT	495
544	GTAAGAGCAACCTTCTGAGAGCAGAAACACAAAGTGTGAGCGCTTTGCATTAAGATTAAGGTG	603	
Db	496	GTAAGAGCAACCTTCTGAGAGCAGAAACACAAAGTGTGAGCGCTTTGCATTAAGATTAAGGTG	555
604	TTGCGTGTGCTTACAGAGATATGACCTCCCTTACTTTGGGGCTCAACGGTCTCTTCGT	663	
Db	556	TTGCGTGTGCTTACAGAGATATGACCTCCCTTACTTTGGGGCTCAACGGTCTCTTCGT	615
664	CATTTAATGGAATCAAGTTCACAGAAATTAACCTGATGGAATTTGATGTTGCCAAGTT	723	
Db	616	CATTTAATGGAATCAAGTTCACAGAAATTAACCTGATGGAATTTGATGTTGCCAAGTT	675
724	TTGGAAGGAGCTTGAAGTACCATGGACAGTTCATTGATTTGTGCATTCCTGTGAGATGT	783	

|||||
Db 676 TTGGAGAGGCTTAACCTACACCATGAGACCAGTTCATTGATTGGCGTCGTGTGGATAT 735
QY 784 GACTATTTGATAGATCAAAAGATATCGGGGGGCAACACGCTGTGAAATCTTATTCGTCAA 843
Db 736 GACTATTTGATAGATCAAAAGATATCGGGGGGCAACACGCTGTGAAATCTTATTCGTCAA 795
QY 844 CATGGGTCCTAGAAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 903
Db 796 CATGGGTCCTAGAAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 855
QY 904 GACGGGCTTACCAAGAGCTGACGCTGTGTCAAGAGCCTATGTCAATGGATATT 963
Db 856 GACTGGCTTACCAAGAGCTGACGCTGTGTCAAGAGCCTATGTCAATGGATATT 915
QY 964 CCTGAGCTAAATGAGCTGACCTGATGAGAGGGTCTCTAAGTTCTCGTAAAGAT 1023
Db 916 CCTGAGCTAAATGAGCTGACCTGATGAGAGGGTCTCTAAGTTCTCGTAAAGAT 975
QY 1024 AATGGTTCAACGAAGATCGGGTGACAAAGCCATAGAGAAGTCAAAATCTGCCAAGAT 1083
Db 976 AATGGTTCAACGAAGATCGGGTGACAAAGCCATAGAGAAGTCAAAATCTGCCAAGAT 1035
QY 1084 AAATCGTCGCAAGAGAGCTCGAGTCCCTTTTCAAGCCAACTGCCACCATCAGCAGCG 1143
Db 1036 AAATCGTCGCAAGAGAGAGCTCGAGTCCCTTTTCAAGCCAACTGCCACCATCAGCAGCG 1095
QY 1144 CTAAACGGAAGAGAGCTCGGATTAACCAAGCAGGAGCTGCGAACAAGAAACAAG 1203
Db 1096 CTAAACGGAAGAGAGAGCTCGGATTAACCAAGCAGGAGCTGCGAACAAGAAACAAG 1155
QY 1204 GCTGGTGAAGAAGAATATCTGTGATGCTTGTGATGATCACTAGACAGTACGAAGCAG 1263
Db 1156 GCTGGTGAAGAAGAATATCTGTGATGCTTGTGATGATCACTAGACAGTACGAAGCAG 1215
QY 1264 CGGTGGGCTGATCACTTCGCTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1323
Db 1216 CGGTGGGCTGATCACTTCGCTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1275
QY 1324 AAAAGTT 1330
Db 1276 AAAAGTT 1282
RESULT 4
ARI52406 1478 bp DNA linear PAT 08-AUG-2001
LOCUS ARI52406
DEFINITION Sequence 7 from patent US 6232527.
ACCESSION ARI52406
VERSION ARI52406.1 GI:15118456
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1478)
AUTHORS Mahajan,P.B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 7 15-MAY-2001.
FEATURES
source location/Qualifiers
1..1478
BASE COUNT 463 a 302 c 365 g 348 t
ORIGIN
Query Match 61.1%; Score 894; DB 6; Length 1478;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

|||||
Db 92 ACGAGATGGGATCAAGAGGTTTGACGAACCTGCTGGCGGACCAATGCCCCAAAGGCATGA 151
QY 140 AGAGCAGAAAGTTGAGAGAGTACTTTCGGCGCAAAATCGCGCTGACGCGCAGATGACA 199
Db 152 AGGAGCAGAAAGTTGAGAGAGTACTTTCGGCGCAAAATCGCGCTGACGCGCAGATGACA 211
QY 200 TATACAGTTCTCTGATGTTGATTTGGAAAGCAGGCAATGGAACCTGCAAAATGAAAGCTG 259
Db 212 TCTACAGTTCTCTGATGTTGATTTGAAAGGACAGGCAATGGAACCTGCAAAATGAAAGCTG 271
QY 260 GTGAAGTACTAGTCAATTTGCAAGAAATGTTCAACCGGACAAATAGATTACTGGAACGG 319
Db 272 GTGAAGTACTAGTCAATTTGCAAGAAATGTTCAACCGGACAAATAGATTACTGGAACGG 331
QY 320 GAATCAAGCCAGTTTATGTTTATGATGCAAGCCTCTGATATGAGAAACAGAGCTTG 379
Db 332 GAATCAAGCCAGTTTATGTTTATGATGCAAGCCTCTGATATGAGAAACAGAGCTTG 391
QY 380 CTAAAGATTAAGTCAAAAGATGATGCAACCAAGATCTGAGGCAAGTGAAGTAG 439
Db 392 CTAAAGATTAAGTCAAAAGATGATGCAACCAAGATCTGAGGCAAGTGAAGTAG 451
QY 440 GAGATTAAGATGCGATTGAAAAATTTGAGCAAGAGAGCTGTAAAGTCAACAAGCACA 499
Db 452 GAGATTAAGATGCGATTGAAAAATTTGAGCAAGAGAGCTGTAAAGTCAACAAGCACA 511
QY 500 ACGAAGATTTGAACGGCTATTAAGACTTATGGGGTTCCCTGTTGTAGAGCAACCTTG 559
Db 512 ACGAAGATTTGAACGGCTATTAAGACTTATGGGGTTCCCTGTTGTAGAGCAACCTTG 571
QY 560 AAGCAAGCAAGAAATGTCACAGCCTTTGATTAACGATTAAGGTGTTGCTTGCCTTCA 619
Db 572 AAGCAAGCAAGAAATGTCACAGCCTTTGATTAACGATTAAGGTGTTGCTTGCCTTCA 631
QY 620 AAGATATGAGTCCCTTACTTTTGGGGCTCCAGGTTCTTTCGATTTTAATGATCAA 679
Db 632 AAGATATGAGTCCCTTACTTTTGGGGCTCCAGGTTCTTTCGATTTTAATGATCAA 691
QY 680 GTTCCAGAAAATACCTGATGATGATTTGATTTGATTTGCAAGGTTTGGAGAGCTTGA 739
Db 692 GTTCCAGAAAATACCTGATGATGATTTGATTTGATTTGCAAGGTTTGGAGAGCTTGA 751
QY 740 TCACCATGAGACAGTTCAATGATTTGTGATCCTGTGTGATGTGACTATTGTGATGA 799
Db 752 TCACCATGAGACAGTTCAATGATTTGTGATCCTGTGTGATGTGACTATTGTGATGA 811
QY 800 TCAAAAGTATCGGGGGGCAAAACAGCTCTGAAACTTATTCGTCACACNTGGGTCCATAGAAA 859
Db 812 TCAAAAGTATCGGGGGGCAAAACAGCTCTGAAACTTATTCGTCACACNTGGGTCCATAGAAA 871
QY 860 GCATCTTGAGATCTTATTAAGACAGATATCAAAATCTGAGAGACTGGCCTTACCAG 919
Db 872 GCATCTTGAGATCTTATTAAGACAGATATCAAAATCTGAGAGACTGGCCTTACCAG 931
QY 920 AAGCTGAGAGCTTTGTTCAAGAGCCTTAATGTACATTGGATTTCTGACCTAAATGCA 979
Db 932 AAGCTGAGAGCTTTGTTCAAGAGCCTTAATGTACATTGGATTTCTGACCTAAATGCA 991
QY 980 CTGCACTGATGAGAGGGGCTCATAGTTTCTGCTGTTAAAGATTAAGGTTTCAACGAG 1039
Db 992 CTGCACTGATGAGAGGGGCTCATAGTTTCTGCTGTTAAAGATTAAGGTTTCAATGAG 1051
QY 1040 ATCGGGTGCAAAAGCCATAGAGAAATGATCAATCTGCCAAGAAATATCGTCGCAAGAA 1099
Db 1052 ATCGGGTGCAAAAGCCATAGAGAAATGATCAATCTGCCAAGAAATATCGTCGCAAGAA 1111
QY 1100 GACTGAGTCTTTTTCAGGCCACTGCACACATCGACCCGCTTAACGGAAGAGA 1159
Db 1112 GACTGAGTCTTTTTCAGGCCACTGCACACATCGACCCGCTTAACGGAAGAGA 1171
QY 1160 CTTGGATTAACCAAGCAGAGCTGCAACAAAGAAAGGCTGTGGAAAGAGA 1219
|||||

Db 1172 CTTGGATATAACAGCAGAGCTGCGACAGCAAGAAAAAAGCGTGGTGGAGAGA 1231

Oy 1220 AATAATCTTGAGCTGATGATCACTACAGTACAGCAAGCGGTGGC 1270
|||||

Db 1232 AATAATCTTGAGCTGATGATCACTACAGTACAGCAAGCGGTGGC 1282

RESULT 5
LOCUS AB021666 1354 bp mRNA linear PLN 25-MAR-2000

DEFINITION *Oryza sativa* OsFEN-1 mRNA for FEN-1, complete cds.

ACCESSION AB021666

VERSION AB021666

KEYWORDS FEN-1; endonuclease.

ORGANISM *Oryza sativa* CDNA to mRNA.

REFERENCE 1 Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; *Oryza*.

REFERENCE 2 (bases 1 to 1354)
Kimura, S., Ueda, T., Hatanaka, M., Takenouchi, M., Hashimoto, J. and Sakauchi, K.
Plant homologue of flap endonuclease-1: molecular cloning, characterization, and evidence of expression in meristematic tissues
Plant Mol. Biol. 42 (3), 415-427 (2000)

REFERENCE 3 (bases 1 to 1354)
Kimura, S., Hashimoto, J. and Sakauchi, K.
Direct Submission
Submitted (21-DEC-1998) Salsuke Kimura, Science University of Tokyo, Dept. of Applied Biological Science; 2641 Yamazaki, Noda, Chiba 278-8510, Japan (E-mail:16498703@ed.noda.sut.ac.jp, Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)
On Apr 17, 1999 this sequence version replaced gi:4062866.

COMMENT On Apr 17, 1999 this sequence version replaced gi:4062866.

FEATURES
source
1..1354
Location/Qualifiers
/organism="Oryza sativa"
/db_xref="taxon:4530"
67..1209
/gene="OsFEN-1"
67..1209
/gene="OsFEN-1"
/note="endonuclease"
/codon_start=1
/product="FEN-1"
/protein_id="BA36171.1"
/db_xref="gi:4587225"
/translation="MGIKGTLKILADNAPKAMKEQFESYFGRILAVDMSIYOFLLVGRGTGMENTINAGEVTGSHLOGMENRTIRLEAGIKPYVYDGGPDLKQELAKRYSKREDATKELTEAVREGDDDAIEFKSRKYTKQHNBECKLLMLKVPYVEACPEAEACALCINDNYAVASDEMDSLTFGARFLRLHMDPSKIPVMEFEVAKVLEELELTMDQFLDLCILSGCDYCSIKIGIGQTLKLRHGSISTELENINFRQVIPEDWPYOBARLFKEPNTLTDIPELKNAPDEBLVFLVKEENFNDRTVTKAIEKIFAKNKSSQGRLESEFFKPVYSTVPLKRDTESEKPTKAVANKTKGAGKKK"
1354
/note="17 A nucleotides"

BASE COUNT 415 a 286 c 331 g 322 t

ORIGIN

Query Match 4.0%; Score 58; DB 8; Length 1354;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 113 TGGCGGACAAATGCGCCCAAGCGCATGAAGGAGCGAGTTCGAGACTACTTCGGCGC 170
|||||

Db 95 TGGCGGACAAATGCGCCCAAGCGCATGAAGGAGCGAGTTCGAGACTACTTCGGCGC 152

RESULT 6
LOCUS AC104713 115907 bp DNA linear HTG 20-DEC-2001

DEFINITION *Oryza sativa* chromosome 5 clone OJ1362G11. *** SEQUENCING IN

ACCESSION AC104713

VERSION AC104713.1 GI:17940800

KEYWORDS HTG; PHASE2.

ORGANISM *Oryza sativa*

REFERENCE 1 Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; *Oryza*.

REFERENCE 2 (bases 1 to 115907)
Chow, T.-Y., Hsing, Y.-I. C., Chen, C.-S., Chen, H.-H., Wu, H.-P., Liu, S.-M., Chao, Y.-T., Chang, S.-J., Chen, T.-R., Chen, Y.-L., Kau, P.-I., Lee, M.-C., Deu, H.-L., Lin, S.-J., Wu, L.-F. and Shaw, J.-F.
Oryza sativa BAC OJ1362G11 genomic sequence
Unpublished

REFERENCE 3 (bases 1 to 115907)
Chow, T.-Y. and Hsing, Y.-I. C.
Direct Submission
Submitted (20-DEC-2001) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
The nucleotide sequence of this BAC clone was generated by combining Monsanto and ASPEC-Taiwan sequencing data.

NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 3719: contig of 3719 bp in length
* 3720 3739: gap of unknown length
* 3740 110847: contig of 107108 bp in length
* 110848 110867: gap of unknown length
* 110868 115907: contig of 5040 bp in length.

FEATURES
source
1..115907
Location/Qualifiers
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="5"
/clone="OJ1362G11"

BASE COUNT 32362 a 24964 c 25806 g 32735 t 40 others

ORIGIN

Query Match 4.0%; Score 58; DB 2; Length 115907;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 113 TGGCGGACAAATGCGCCCAAGCGCATGAAGGAGCGAGTTCGAGACTACTTCGGCGC 170
|||||

Db 8213 TGGCGGACAAATGCGCCCAAGCGCATGAAGGAGCGAGTTCGAGACTACTTCGGCGC 8156

RESULT 7
LOCUS HSJ104017 105137 bp DNA linear ERI 15-FEB-2000

DEFINITION Human DNA sequence from clone RPL-104017 on chromosome 9q16.1-16.3. Contains STRs and GSSs, complete sequence.

ACCESSION AL080316

VERSION AL080316.8 GI:5650657

KEYWORDS HTG.

ORGANISM human.

REFERENCE 1 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia; Eutheria: Primates: Catarrhini: Homiidae: Homo.

REFERENCE 2 (bases 1 to 105137)
Tracey, A.
Direct Submission
Submitted (14-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone

COMMENT

requests: clonerequest@sanger.ac.uk
On Jul 29, 1999 this sequence version replaced g1:5566542.
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep RPI-104017 is from the library RPI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>

This sequence is the entire insert of clone RPI-104017. This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/MGP/Ch6>.

FEATURES

source
location/Qualifiers
1..105137
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="q16.1-16.3"
/clone_lib="RPI-1"
/clone="RPI-104017"
complement(372..811)
/note="match: GSS: Em:A0006718"
1037..1080
/note="22 copies 2 mer ta 88% conserved"
1153..1456
/note="AluSx repeat: matches 1..301 of consensus"
2365..2702
/note="L2 repeat: matches 2363..2750 of consensus"
3579..3733
/note="L2 repeat: matches 2041..2201 of consensus"
4961..5272
/note="AluY repeat: matches 1..311 of consensus"
5545..5648
/note="52 copies 2 mer ta 71% conserved"
7388..7668
/note="AluJo repeat: matches 2..283 of consensus"
7871..8066
/note="MER58A repeat: matches 10..224 of consensus"
complement(10286..10857)
/note="match: GSS: Em:B82911"
11075..11346
/note="L1ME3 repeat: matches 5657..5940 of consensus"
complement(13059..13460)
/note="match: STS: Em:G21604"
14002..14199
/note="MER58C repeat: matches 3..89 of consensus"
17080..17249
/note="L1MC1 repeat: matches 6165..6323 of consensus"
17425..17538
/note="MIR repeat: matches 131..245 of consensus"
complement(17580..17784)
/note="match: STS: Em:G43022"
18332..18525
/note="AluJo repeat: matches 87..295 of consensus"
18919..19338
/note="L2 repeat: matches 1868..2326 of consensus"
19344..19733

/note="MIR1B repeat: matches 12..389 of consensus"
20175..20286
/note="56 copies 2 mer at 64% conserved"
20512..20976
/note="L2 repeat: matches 2264..2750 of consensus"
21113..21217
/note="L2 repeat: matches 1233..1341 of consensus"
21331..21517
/note="L1MC repeat: matches 1768..1953 of consensus"
21882..22020
/note="L1ME repeat: matches 419..560 of consensus"
22615..22710
/note="48 copies 2 mer tt 62% conserved"
23360..23417
/note="29 copies 2 mer tt 69% conserved"
24769..25327
/note="HERV16 repeat: matches 6..548 of consensus"
25328..26196
/note="MER6 repeat: matches 1..864 of consensus"
26197..26818
/note="HERV16 repeat: matches 548..1219 of consensus"
26964..27296
/note="JTR16A repeat: matches 95..450 of consensus"
27297..27643
/note="HERV1 repeat: matches 3587..3946 of consensus"
28215..28338
/note="MST1 repeat: matches 284..422 of consensus"
28589..29009
/note="match: GSS: Em:A0308663"
28617..28812
/note="MST1 repeat: matches 1..178 of consensus"
29010..29383
/note="MIR1B repeat: matches 1..388 of consensus"
30078..30297
/note="JTR16A repeat: matches 15..229 of consensus"
32881..33162
/note="L1MC5 repeat: matches 7649..7926 of consensus"
33433..33616
/note="MIR repeat: matches 40..260 of consensus"
complement(33617..33834)
/note="match: GSS: Em:A0240281"
33899..34036
/note="MIR1 repeat: matches 269..404 of consensus"
37901..38084
/note="AluS9/X repeat: matches 83..264 of consensus"
38096..38119
/note="L2 copies 2 mer ac 95% conserved"
38855..38952
/note="MIR repeat: matches 85..184 of consensus"
39820..40101
/note="AluY repeat: matches 1..299 of consensus"
40284..41588
/note="L1PA7 repeat: matches 2138..3447 of consensus"
41588..44294
/note="L1PA7 repeat: matches 3441..6142 of consensus"
45641..46728
/note="L1ME1 repeat: matches 4994..6109 of consensus"
46737..47808
/note="TIGOR1 repeat: matches 1350..2418 of consensus"
47815..48092
/note="Charliefb repeat: matches 70..386 of consensus"
48179..48335
/note="L1PB1 repeat: matches 5999..6155 of consensus"
48338..48369
/note="16 copies 2 mer aa 93% conserved"
48370..48882
/note="L1M4 repeat: matches 4272..4790 of consensus"
48927..49175
/note="L1MC repeat: matches 2146..2412 of consensus"
49180..49590
/note="L1M3 repeat: matches 7310..7739 of consensus"
49630..49717
/note="L1MC4 repeat: matches 7219..7313 of consensus"


```

repeat_region      49732..50676
                    /note="LMC repeat: matches 1514..2152 of consensus"
repeat_region      50716..50836
                    /note="LMC repeat: matches 1287..1413 of consensus"
repeat_region      50849..51871
                    /note="LMC1 repeat: matches 5287..6316 of consensus"
repeat_region      51874..52239
                    /note="LMC repeat: matches 929..1302 of consensus"
repeat_region      52268..52519
                    /note="LMC repeat: matches 523..780 of consensus"
repeat_region      53098..53206
                    /note="Alu repeat: matches 197..305 of consensus"
repeat_region      53208..53388
                    /note="MIR repeat: matches 41..231 of consensus"
repeat_region      53414..53830
                    /note="L1P8 repeat: matches 5746..6163 of consensus"
repeat_region      54168..54264
                    /note="MIR repeat: matches 2..111 of consensus"
misc_feature       complement(54484..55028)
                    /note="match: GSS: Em:A0318586"
repeat_region      55956..56266
                    /note="AluX repeat: matches 1..312 of consensus"
repeat_region      56471..56682
                    /note="MIR repeat: matches 6..232 of consensus"
repeat_region      56744..56947
                    /note="L2 repeat: matches 2246..2448 of consensus"
repeat_region      57428..57463
                    /note="L18 copies 2 mer tt 100% conserved"
misc_feature       complement(57599..58119)
                    /note="match: GSS: Em:A0346372"
repeat_region      58770..59081
                    /note="Alu repeat: matches 1..309 of consensus"
repeat_region      59703..60000
                    /note="Alu repeat: matches 1..294 of consensus"
repeat_region      60141..60194
                    /note="L2 repeat: matches 2652..2706 of consensus"
repeat_region      61142..61188
                    /note="L2 repeat: matches 2652..2698 of consensus"
repeat_region      61806..61910
                    /note="L2 repeat: matches 2652..2698 of consensus"

Query Match
Best Local Similarity 100.0%; Pred. No. 6.8e-13; Length 105137;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1421 CCTTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 57474 CCTTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 57432

RESULT 8
BC022967      2871 bp      mRNA      linear      PRI 07-FEB-2002
LOCUS      Homo sapiens, clone MGC:30000 IMAGE:5111152, mRNA, complete cds.
DEFINITION      BC022967
ACCESSION      BC022967.1 GI:18605802
VERSION      MGC.
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2871)
AUTHORS      Strausberg,R.
TITLE      Direct Submission
JOURNAL      Submitted (04-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@email.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

```

```

FEATURES
    source
        DNA Sequencing by: Baylor College of Medicine Human Genome
        Sequencing Center
        Center code: BCM-HGSC
        Web site: http://www.hgsc.bcm.tmc.edu/cdna/
        Contact: amg@bcm.tmc.edu
        Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
        Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
        Richards, S., Gibbs, R.A.
        Clone distribution: MGC clone distribution information can be found
        through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
        Series: IRAX Plate: 42 Row: m Column: 13
        This clone was selected for full length sequencing because it
        passed the following selection criteria: Hexamer frequency ORF
        analysis, Genomescan gene prediction.
        location/Qualifiers
            1..2871
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="MGC:30000 IMAGE:5111152"
            /issue_type="Cervix, Carcinoma"
            /clone_lib="NIH_MGC_12"
            /lab_host="DH10B"
            /note="Vector: PCMV-SPORT6"
            7..987
            /product="Unknown (protein for MGC:30000)"
            /protein_id="AAH22867.1"
            /db_xref="GI:18603803"
            /translation="MKRRAGIGSMRSVVGFLSGRLGDLPLTODFORRLRGCRNL
            YKKDLGHRGCVNAIEFSNNGQMLVSGDDRRVLLMHMEQAHISRVPIOLKGHS
            NIECLAFNSGNTKFEVSGNDEQVILHDVSTLDPFHEDAVYGLSVSPVNDNFAS
            SSDDGVLLIDIRRESPHGEPFLANPSAFHVMNPVPEPLRLAANSKEGVLMDIR
            KQSSILRYGGLSLQASMSVRFNSNGTGLARLRPLPYLDIRLPPYFQPNQGT
            FNSCTMKSCCFGRDQDHNMKRRQANSFQFLSLTRNGDADVDKLAV"

BASE COUNT      802 a      569 c      635 g      865 t

Query Match
Best Local Similarity 100.0%; Pred. No. 3e-11; Length 2871;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 TTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2829 TTGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2868

RESULT 9
AC084072      89181 bp      DNA      linear      HTG 05-FEB-2002
LOCUS      Mus musculus clone rp22-51112 strain 12956/SVEV7ac, WORKING DRAFT
DEFINITION      AC084072
SEQUENCE      17 unordered pieces.
ACCESSION      AC084072
VERSION      AC084072.6 GI:18497111
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 89181)
AUTHORS      Zhu,H., Zullo,S.B. and Roe,B.A.
TITLE      Mus musculus BAC Clone rp22-51112
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 89181)
AUTHORS      Wu,H., Hu,P., Yang,L., Prescott,A., Zullo,S.J. and Roe,B.A.
TITLE      Direct Submission
JOURNAL      Submitted (12-OCT-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REMARK      On Feb 5, 2002 this sequence version replaced gi:18390228.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma

```

Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2251: contig of 2251 bp in length
* 2252 2351: gap of unknown length
* 2352 4397: contig of 2046 bp in length
* 4398 4497: gap of unknown length
* 4498 6900: contig of 2403 bp in length
* 6901 7000: gap of unknown length
* 7001 9132: contig of 2132 bp in length
* 9133 9232: gap of unknown length
* 9233 12520: contig of 3288 bp in length
* 12521 12620: gap of unknown length
* 12621 14871: contig of 2251 bp in length
* 14872 14971: gap of unknown length
* 14972 17428: contig of 2457 bp in length
* 17429 17528: gap of unknown length
* 17529 22646: contig of 5118 bp in length
* 22647 22746: gap of unknown length
* 22747 26265: contig of 3519 bp in length
* 26266 26365: gap of unknown length
* 26366 30020: contig of 3655 bp in length
* 30021 30120: gap of unknown length
* 30121 34472: contig of 4352 bp in length
* 34473 34572: gap of unknown length
* 34573 40233: contig of 5661 bp in length
* 40234 40333: gap of unknown length
* 40334 45943: contig of 5610 bp in length
* 45944 46043: gap of unknown length
* 46044 51644: contig of 5601 bp in length
* 51645 51744: gap of unknown length
* 51745 60274: contig of 8530 bp in length
* 60275 60374: gap of unknown length
* 60375 71451: contig of 11077 bp in length
* 71452 71551: gap of unknown length
* 71552 89181: contig of 17630 bp in length.
* Location/Qualifiers
1. 89181

FEATURES

source

/organism="Mus musculus"
/strain="129S6/SVEVtac"
/db_xref="taxon:10090"
/clone="rp22-51112"

BASE COUNT 26250 a 17373 c 16270 g 27668 t 1620 others

ORIGIN

Query Match 2.7%; Score 40; DB 2; Length 89181;
Best Local Similarity 100.0%; Pred. No. 3.5e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 TTGAAGTAAAAA 1463

Db 60839 TTGAAGTAAAAA 60878

RESULT 10

AX185400/c 288 bp DNA linear PAT 06-AUG-2001

LOCUS AX185400 Sequence 1095 from Patent WO0142467.

ACCESSION AX185400

VERSION AX185400.1 GI:15136793

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (Bases 1 to 288)

AUTHORS Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 1095 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES

source

1. 288
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT

88 a 49 c 42 g 107 t 2 others

ORIGIN

Query Match 2.7%; Score 39; DB 6; Length 288;
Best Local Similarity 100.0%; Pred. No. 9.9e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1423 TTGAAGTAAAAA 1461

Db 66 TTGAAGTAAAAA 28

RESULT 11

AX186826/c 305 bp DNA linear PAT 06-AUG-2001

LOCUS AX186826 Sequence 2521 from Patent WO0142467.

DEFINITION AX186826

ACCESSION AX186826

VERSION AX186826.1 GI:15138268

KEYWORDS

SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 305)
Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 2521 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES

source

Location/Qualifiers

1. 305

/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 96 a 45 c 51 g 113 t

ORIGIN

Query Match 2.7%; Score 39; DB 6; Length 305;
Best Local Similarity 100.0%; Pred. No. 9.9e-11;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1423 TTGAAGTAAAAA 1461

Db 44 TTGAAGTAAAAA 6

RESULT 12

PFMA14P4 22448 bp DNA linear HTG 11-AUG-1999

LOCUS PFMA14P4 Plasmodium falciparum chromosome 4 strain 3D7, *** SEQUENCING IN

DEFINITION PROGRESS ***, in unordered pieces.

ACCESSION AU035477 GI:5731932

VERSION AU035477.5

KEYWORDS HTG: HTGS_PHASE1.

SOURCE malaria parasite P. falciparum.

ORGANISM

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 22448)

AUTHORS Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M.
and Barrett,I.B.

TITLE Direct Submission

JOURNAL

Submitted (24-FEB-1999) P. falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK

COMMENT

On Aug 12, 1999 this sequence version replaced q1:5531400.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum. IMPORTANT: This
sequence is unfinished and does not necessarily represent the
correct sequence. Work on the sequence is in progress and the
release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc.
Order of segments is not known: 800 n's separate segments.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES

Location/Qualifiers

1..22448
/organism="Plasmodium falciparum"
/strain="3D7"
/db_xref="taxon:5833"
/chromosome="4"

BASE COUNT 80792 a 23258 c 21957 g 75240 t 23201 others

ORIGIN

Query Match 2.7%: Score 39; DB 2; Length 22448;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1425 TGAAGTAAAAA
Db 40702 TGAAGTAAAAA
1463
40740

RESULT 13

BC009510 659 bp mRNA linear PRI 12-JUN-2001
LOCUS Homo sapiens, Similar to RIKEN cDNA 3110001D03 gene, clone MGC:4730
DEFINITION IMAGE:353614, mRNA, complete cds.
ACCESSION BC009510
VERSION BC009510.1 GI:14550519
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
contact: <http://www.systemsbio.org>
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

FEATURES

Location/Qualifiers

1..659
/organism="Homo sapiens"
/db_xref="taxon:9606"

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 11 Row: 1 Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

CDS

/clone="MGC:4730 IMAGE:353614"
/tissue_type="Lung, small cell carcinoma"
/clone_id="NIH-MGC-7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
43..393
/codon_start=1
/product="Similar to RIKEN cDNA 3110001D03 gene"
/protein_id="AAH09510.1"
/db_xref="GI:14550520"
/translation="MGSRLSQPFESYITAPGTAAPAKPAPATPGAPTPAERRL
KTCMSCPVLSGLGIMGAGVYVWAKPKMKGYPPSPMTTQVIGLSNGLATWGI
VWADPKGRKRYRV"

BASE COUNT 178 a 172 c 174 g 135 t

ORIGIN

Query Match 2.6%: Score 38; DB 9; Length 659;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAA
Db 615 GAAGTAAAAA
1463
652

RESULT 14

AX011676 691 bp DNA linear PAT 06-SEP-2000
LOCUS Sequence 72 from Patent W0955858.
DEFINITION AX011676
ACCESSION AX011676
VERSION AX011676.1 GI:998200
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 691)
Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenchal, A. and
Pillarsky, C.
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue
JOURNAL Patent: WO 9955858-A 72 04-NOV-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES Location/Qualifiers
1..691
source
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 192 a 181 c 174 g 144 t

ORIGIN

Query Match 2.6%: Score 38; DB 6; Length 691;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAA
Db 631 GAAGTAAAAA
1463
668

RESULT 15

HSMB02710 2228 bp mRNA linear PRI 28-JUN-2000
LOCUS Homo sapiens mRNA, cDNA DKFp57C146 (from clone DKFp547C146).
DEFINITION AL359947
ACCESSION AL359947
VERSION AL359947.1 GI:8977907
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2228)
Blum, H., Bauersachs, S., Mewes, H.W., Well, B. and Wilmann, S.

TITLE Direct Submission
JOURNAL Submitted (15-JUN-2000) MIPS, Am Klopferstritz 18a, D-82152 Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Munich/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp547C146) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES
source
1. .2228
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="547 (synonym: hfb1). Vector psport1; host DH10B; sites NotI + SalI"
/dev_stage="fetal"
/tissue_type="brain"
2184

polyA_site
BASE COUNT 694 a 403 c 386 g 745 t
ORIGIN

Query Match 2.6%; Score 38; DB 9; Length 2228;
Best Local Similarity 100.0%; Pred. No. 4.1e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1426 GAAGTAAAAA
Db 2184 GAAGTAAAAA

RESULT 16
AB042653 2240 bp mRNA linear MAM 23-AUG-2001
LOCUS Bos taurus mRNA for zona pellucida glycoprotein ZPA, complete cds.
DEFINITION AB042653
ACCESSION AB042653.1 GI:12583656
VERSION
KEYWORDS
SOURCE Bos taurus cDNA to mRNA, clone_lib:lambda ZAP bovine ovary cDNA (stratagene).
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 2240)
Yonezawa, N., Fukui, N., Kuno, M., Shinoda, M., Goko, S., Mitsui, S. and Nakano, M.
Molecular cloning of bovine zona pellucida glycoproteins ZPA and ZPB and analysis for sperm-binding component of the zona
European Journal of Biochemistry / FEBS. 268 (12), 3587-3594 (2001)
21315847
2 (bases 1 to 2240)
Nakano, M. and Yonezawa, N.
Direct Submission
Submitted (15-MAY-2000) Minoru Nakano, Chiba University, Department of Chemistry, Inage-ku, 1-33 Yayoi-cho, Chiba 263-8522, Japan (E-mail: mmakano@sci.chem.s.chiba-u.ac.jp, Tel:81-43-290-2794, Fax:81-43-290-2874)

FEATURES
source
1. .2240
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="lambda ZAP bovine ovary cDNA (stratagene)"
30. .2171
/gene="ZPA"
30. .2171
/gene="ZPA"
/note="Vector: pOTB7"
/codon_start=1
/product="zona pellucida glycoprotein ZPA"

/protein_id="BAB21482.1"
/db_xref="GI:12583657"
/translation="MACRODSDGRSPSSWFEADRSFPLSTLTYSNIDVNOIDPA
FEGIVATCYENKMYVEFKTLTNKQHSYVDSGLKMNICNYVLDPEKTLTKAPYESC
TKRVLGQHMVTTEENDNTAHRKQTLTVHSCVPWQAGRHOHSGSTICKRSPF
HEFPLADDTAGPKPQMGWTVYVDGEAQLVQLAEALDQYINLLIENOKSIOVLRH
ATGVTHSQGNSHLYMPLKTHVSPGQITLSSRLICADPPVCNMTHTLTPSP
GRKVSFENKNINAVNOLNNGIYVKEINGLRLHFSKTLTKFSEKCLPYOYLSL
KLTFYOLETVSMVYTPCEVESTYISGELCTODGFMDEVYRHQTKPALNDTLR
VDSGSCQPTKAPROGILKYPHPIINGCQTRKFPNGKVIYENEIHALMADLPSTISR
DSERHRTVRCYSSSNMLTNVBSLPPVAVSRGPAITLQYVPPNSYLOPGDND
YPVVRYLQPTYLEVRLNRDIPNKLIDCQKATSTYDPAISLQNNIYDGCYVND
NHRTHFHVGSVAVPNHYORFAKTEFVSDPAFSLVIFCSALICDOLSNFEL
CSASCLVSSRSRRATGATEEEKMIVSLPILLLSDGSSFRDAVDSKGHTSGYAAK
TWAVVALAGVAVATLSLSTLRKKRIYVNLN"

BASE COUNT 635 a 536 c 485 g 584 t
ORIGIN

Query Match 2.6%; Score 38; DB 4; Length 2240;
Best Local Similarity 100.0%; Pred. No. 4.1e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1426 GAAGTAAAAA
Db 2193 GAAGTAAAAA

RESULT 17
BC001537 2291 bp mRNA linear PRI 22-OCT-2001
LOCUS Homo sapiens, clone MGC:786 IMAGE:2987938, mRNA, complete cds.
DEFINITION BC001537
ACCESSION BC001537.1 GI:16306716
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2291)
Strausberg, R.
Direct Submission
Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cga@gsf-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan

REMARK
COMMENT
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 3 Row: 3 Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657474.

FEATURES
source
1. .2291
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:786 IMAGE:2987938"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_15"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
51. .1820
/codon_start=1

/product="Unknown (protein for MG:786)"
 /protein_id="AAH01537.1"
 /db_xref="GI:16306717"
 /translation="MAADGDSLYPIAVLIDELRNEVQRLNSIKLSTYALGV
 ETRRELLPFLDTITDEDEVLALAEOLGFTTLVGSPVHCLPLESLAYEET
 VVRDAVSLRLAISHESPSDLEAFHVLVRLAGDWFSTRSACGLSYCYRVS
 AVKALROYFRNLCSDDTPMVRRAASLGEFAKLELDNKSEIIPMSNLASPEOD
 SVRLAVACVNIQOLLPOEDIEALVMTLROAEDKSWRVYVADKETEOKAVGP
 ETTKIDLVAFONLMDCEAEVRAASHKVEKPCNLSADCEVNIQIIIPCKEY
 SPANOHVSALASVIMGLSPITGKNTIEHLPLFLADLKECEPEVRLNTISLDCVN
 EYIGTRQSSQSLIPAIVELAEDAKRWVLAITEVPLVLAGDGLVEFEDEKLSICMAV
 LVDHVAIREATSNLKKLVEKFKEMAHATIPKVLVMSGDPNLTNHRMTLFCINVL
 SEVCGDDITTKHMLPVLVLMAGDPPVAVNRFVNAKSLOKIGPILDNSTLOSEVRLKLEK
 LTQDDVDVKKFAEALTIVLSIA"

BASE COUNT 490 a 682 c 642 g 477 t

Query Match 2.6%; Score 38; DB 9; Length 2291;
 Best Local Similarity 100.0%; Pred. No. 4.1e-10;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GAAGTAAAAA... 1463
 Db 2252 GAAGTAAAAA... 2289

RESULT 18
 LOCUS AY069109 3315 bp mRNA linear INV 17-DEC-2001
 DEFINITION Drosophila melanogaster GH12580 full length cDNA.
 ACCESSION AY069109
 VERSION AY069109.1 GI:17861553
 KEYWORDS FLI_CDNA.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 3315)
 Stapleton, M., Brockslein, P., Hong, L., Aghayani, A., Carlson, J.,
 Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
 Gonzalez, M., Martin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
 Nuno, J., Paclob, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,
 Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.
 Direct Submission
 Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA

COMMENT
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy, presence of a polyA tail and configity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcription of unspliced precursor RNAs, and
 reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our Web site
 (<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES
 SOURCE
 1..3315
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /strain="y; cn bw sp"
 /db_xref="taxon:7227"
 /map="86C2-86C2"

gene
 1..3315
 /gene="Skeletor"
 /note="alignment with genomic scaffold AE003688"
 /db_xref="FLYBASE:FBgn0020517"
 1933..3054
 /gene="Skeletor"
 /note="Longest ORF"
 /codon_start=1
 /db_xref="FLYBASE:FBgn0020517"
 /product="GH12580p"
 /protein_id="AAL39254.1"
 /db_xref="GI:17861554"
 /translation="MRPVRVQGRKPDPTVYVESGFKPIVITDGTGVLPKEITIDYVHR
 RBPQGTIDVMEITDITFLAQQGSGSTOSFEFMTPTSPDSTNATKVLRYVVKVSP
 TASALRTPSAALEHALPSASELTKPTLDELFADLEBELEMPFVADVESLEET
 KDAVYTTINIPRNTTKRPPDLLLEDLFGDEEELAYADELEMDRVAAMARIETY
 YLPDNRKIPDTRVPSGALYTFDGKSVDSLVLPKLDAPDANVHONAOYGLPL
 EOLVRTYPRGVYRGELPQERGTPEOPVSEVSHPAPEFRTTPVVFSSSGSTIIYPRSS
 STGASVTVSSSSPLSSSLRPSTIKIQLKPEGRRA"

BASE COUNT 880 a 992 c 812 g 631 t

Query Match 2.6%; Score 38; DB 3; Length 3315;
 Best Local Similarity 100.0%; Pred. No. 4.1e-10;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GAAGTAAAAA... 1463
 Db 3274 GAAGTAAAAA... 3311

RESULT 19
 LOCUS AC105102/c 67632 bp DNA linear HTG 23-DEC-2001
 DEFINITION Homo sapiens chromosome 17 clone RP11-316K20 map 17, LOW-PASS
 SEQUENCE SAMPLING.
 ACCESSION AC105102
 VERSION AC105102.1 GI:17977598
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 67632)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 17, clone RP11-316K20
 Unpublished
 2 (bases 1 to 67632)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguski, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., DeArillano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardina, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
 Hages, B., Headford, A., Horton, L., Hulme, W., Illiev, I., Johnson, R.,
 Jones, C., Kamt, A., Karatas, A., Kells, C., Larocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McGowan, P., McKernan, R., McPherson, R., Meldrum, J.,
 Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (23-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRB

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L22879

Center clone name: 316_K_20

* NOTE: This record contains 80 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 748: contig of 748 bp in length
749 848: gap of 100 bp
849 1565: contig of 717 bp in length
1566 1665: gap of 100 bp
1666 2391: contig of 726 bp in length
2392 2491: gap of 100 bp
2492 3242: contig of 751 bp in length
3243 3342: gap of 100 bp
3343 4107: contig of 765 bp in length
4108 4207: gap of 100 bp
4208 4975: contig of 768 bp in length
4976 5075: gap of 100 bp
5076 5833: contig of 758 bp in length
5834 5933: gap of 100 bp
5934 6686: contig of 753 bp in length
6687 6786: gap of 100 bp
6787 7545: contig of 759 bp in length
7546 7645: gap of 100 bp
7646 8374: contig of 729 bp in length
8375 8474: gap of 100 bp
8475 9218: contig of 744 bp in length
9219 9318: gap of 100 bp
9319 10059: contig of 741 bp in length
10060 10159: gap of 100 bp
10160 10901: contig of 742 bp in length
10902 11001: gap of 100 bp
11002 11748: contig of 747 bp in length
11749 11848: gap of 100 bp
11849 12609: contig of 761 bp in length
12610 12709: gap of 100 bp
12710 13472: contig of 763 bp in length
13473 13572: gap of 100 bp
13573 14310: contig of 738 bp in length
14311 14410: gap of 100 bp
14411 15163: contig of 753 bp in length
15164 15263: gap of 100 bp
15264 16023: contig of 760 bp in length
16024 16123: gap of 100 bp
16124 16872: contig of 749 bp in length
16873 16972: gap of 100 bp
16973 17717: contig of 745 bp in length
17718 17817: gap of 100 bp
17818 18561: contig of 744 bp in length
18562 18661: gap of 100 bp
18662 19402: contig of 741 bp in length
19403 19502: gap of 100 bp
19503 20217: contig of 715 bp in length
20218 20317: gap of 100 bp
20319 21054: contig of 737 bp in length
21055 21154: gap of 100 bp

21155 21911: contig of 757 bp in length
21912 22011: gap of 100 bp
22012 22770: contig of 759 bp in length
22771 22870: gap of 100 bp
22871 23627: contig of 757 bp in length
23628 23727: gap of 100 bp
23728 24480: contig of 753 bp in length
24481 24580: gap of 100 bp
24581 25327: contig of 747 bp in length
25328 25427: gap of 100 bp
25428 26179: contig of 752 bp in length
26180 26279: gap of 100 bp
26280 27038: contig of 759 bp in length
27039 27138: gap of 100 bp
27139 27886: contig of 748 bp in length
27887 27986: gap of 100 bp
27987 28738: contig of 752 bp in length
28739 28838: gap of 100 bp
28839 29582: contig of 744 bp in length
29583 29682: gap of 100 bp
29683 30427: contig of 745 bp in length
30428 30527: gap of 100 bp
30528 31245: contig of 718 bp in length
31246 31345: gap of 100 bp
31346 32107: contig of 762 bp in length
32108 32207: gap of 100 bp
32208 32972: contig of 765 bp in length
32973 33072: gap of 100 bp
33073 33828: contig of 756 bp in length
33829 33928: gap of 100 bp
33929 34693: contig of 765 bp in length
34694 34793: gap of 100 bp
34794 35535: contig of 742 bp in length
35536 35635: gap of 100 bp
35636 36381: contig of 746 bp in length
36382 36481: gap of 100 bp
36482 37177: contig of 696 bp in length
37178 37277: gap of 100 bp
37278 38004: contig of 727 bp in length
38005 38104: gap of 100 bp
38105 38821: contig of 717 bp in length
38822 38921: gap of 100 bp
38922 39675: contig of 754 bp in length
39676 39775: gap of 100 bp
39776 40525: contig of 750 bp in length
40526 40625: gap of 100 bp
40626 41380: contig of 755 bp in length
41381 41480: gap of 100 bp
41481 42244: contig of 764 bp in length
42245 42344: gap of 100 bp
42345 43098: contig of 754 bp in length
43099 43198: gap of 100 bp
43199 43944: contig of 746 bp in length
43945 44044: gap of 100 bp
44045 44800: contig of 756 bp in length
44801 44900: gap of 100 bp
44901 45660: contig of 760 bp in length
45661 45760: gap of 100 bp
45761 46514: contig of 754 bp in length
46515 46614: gap of 100 bp
46615 47355: contig of 741 bp in length
47356 47455: gap of 100 bp
47456 48214: contig of 759 bp in length
48215 48314: gap of 100 bp
48315 49059: contig of 745 bp in length
49060 49159: gap of 100 bp
49160 49919: contig of 760 bp in length
49920 50019: gap of 100 bp
50020 50725: contig of 706 bp in length
50726 50825: gap of 100 bp
50826 51581: contig of 756 bp in length
51582 51681: gap of 100 bp
51682 52420: contig of 739 bp in length

```

* 52421 52520: gap of 100 bp
* 52521 53253: contig of 733 bp in length
* 53254 53353: gap of 100 bp
* 53354 54092: contig of 739 bp in length
* 54093 54192: gap of 100 bp
* 54193 54927: contig of 735 bp in length
* 54928 55027: gap of 100 bp
* 55028 55724: contig of 697 bp in length
* 55725 55824: gap of 100 bp
* 55825 56595: contig of 771 bp in length
* 56596 56695: gap of 100 bp
* 56696 57447: contig of 752 bp in length
* 57448 57547: gap of 100 bp
* 57548 58290: contig of 743 bp in length

```

Query Match 2.6%: Score 38; DB 2; Length 67632;
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||||||||||||||||||||||||||||||||||||||
 Db 35063 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 35026

RESULT 20
 PPMALIP2/c 98734 bp DNA linear HTG 11-AUG-1999
 LOCUS Plasmidium falciparum chromosome 1 strain 3D7, *** SEQUENCING IN
 DEFINITION PROGRESS ***, in unordered pieces.
 ACCESSION AL031745
 VERSION AL031745.7 GI:5731893
 KEYWORDS HTG; HTGS; PHASE1.
 SOURCE malarial parasite P. falciparum.
 ORGANISM Plasmidium falciparum
 Eukaryotes; Alveolates; Apicomplexa; Haemosporidia; Plasmodium.
 REFERENCE 1 (bases 1 to 98734)
 AUTHORS Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M.
 TITLE Direct Submission
 JOURNAL Submitted (24-SEP-1998) P. falciparum Genome Sequencing Consortium,
 The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
 CB10 1SA, UK
 COMMENT On Aug 12, 1999 this sequence version replaced gi:5706497.
 For more information about this sequence or the Malaria Project,
 see <http://www.sanger.ac.uk/Projects/P-falciparum>. IMPORTANT: This
 sequence is unfinished and does not necessarily represent the
 correct sequence. Work on the sequence is in progress and the
 release of this data is based on the understanding that the
 sequence may change as work continues. The sequence may be
 contaminated with foreign sequence from E.coli, yeast, vector,
 phage etc.
 * NOTE: This is a 'working draft' sequence.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES

source

1. .98734
 /organism="Plasmidium falciparum"
 /strain="3D7"
 /db_xref="taxon:5833"
 /chromosome="1"
 BASE COUNT 37227 a 8619 c 9384 g 35504 t 8000 others
 ORIGIN

Query Match 2.6%: Score 38; DB 2; Length 98734;
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||||||||||||||||||||||||||||||||||||||
 Db 16502 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 16465

RESULT 21
 AL354982/c 119077 bp DNA linear PRI 16-OCT-2001
 LOCUS Human DNA sequence from clone RP11-473H13 on chromosome 9, complete
 DEFINITION sequence.
 ACCESSION AL354982
 VERSION AL354982.12 GI:16214577
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 119077)
 AUTHORS Clark, S.
 JOURNAL Direct Submission
 Submitted (16-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk clone
 requests: clonerequests@sanger.ac.uk

COMMENT

On Oct 17, 2001 this sequence version replaced gi:15131224.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em., EMBL; Sw.,
 SWISSPROT; Tr., TrEMBL; Wp., WormPEP; information on the WormPEP
 database can be found at
<http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr9>
 RP11-473H13 is from the library RCT-11.2 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-473H13. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true left end of clone RP11-410X21 is at 117078 in this
 sequence. The true right end of clone RP11-427J11 is at 2000 in
 this sequence.

FEATURES

source

1. .119077
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-473H13"
 /clone_1lb="RPCT-11.2"
 BASE COUNT 36688 a 22359 g 37984 t
 ORIGIN

Query Match 2.6%: Score 38; DB 9; Length 119077;
 Best Local Similarity 100.0%; Pred. No. 4.9e-10;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||||||||||||||||||||||||||||||||||||||
 Db 83006 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 82969

RESULT 22
 HS199L16/c

LOCUS H5199L16 134963 bp DNA linear PRI 23-NOV-1999
 DEFINITION Human DNA sequence from clone 199L16 on chromosome Xq22.1-22.3
 Contains a 40S Ribosomal Protein S29 Like (pseudor) gene, an EST,
 STSs, GSSs and tcta and tatg repeat polymorphisms, complete
 sequence.
 ACCESSION AL022151.1 GI:3790137
 VERSION AL022151.1
 KEYWORDS HTG: 40S Ribosomal Protein S29; tatg repeat polymorphisms; tcta
 repeat polymorphisms.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 134963)
 AUTHORS Errington,H.
 TITLE Direct Submission
 JOURNAL Submitted (14-OCT-1998) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT requests: clonequest@sanger.ac.uk
 On Oct 26, 1998 this sequence version replaced gi:3646052.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence is the entire insert of clone 199L16. The true right
 end of clone 468H17 (AL031114) is at 26354 in this sequence.
 This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 This sequence was generated from part of bacterial clone contigs of
 human chromosome X, constructed by the Sanger Centre Chromosome X
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/ChrX
 199L16 is from the library RPCI1 constructed at the Roswell Park
 Cancer Institute by the group of Pieter de Jong. For further
 details see http://hacpac.med.buffalo.edu/VECTOR: pcrPAC2.
 FEATURES
 source
 1..134963
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="q22.1-22.3"
 /clone="RPI-199L16"
 /clone_lib="RPCI-1"
 1..237
 /note="AluX repeat: matches 1..237 of consensus"
 239..5892
 /note="L1M1 repeat: matches -1390..4370 of consensus"
 6073..6262
 /note="L1MC/D repeat: matches 5493..5700 of consensus"
 6264..6351
 /note="AluSg/X repeat: matches 210..297 of consensus"
 6355..8296
 /note="L1MB1 repeat: matches 3986..5896 of consensus"
 8291..8555
 /note="L1PB repeat: matches 257..523 of consensus"
 8559..8584
 /note="L1PB repeat: matches 257..523 of consensus"
 8585..8876
 /note="L1M4C repeat: matches 1990..2015 of consensus"
 8877..9023
 /note="L1PA8 repeat: matches 5874..6162 of consensus"
 9024..9128
 /note="L1M4C repeat: matches 1869..1991 of consensus"
 9122..9374
 /note="L1M4C repeat: matches 1869..1991 of consensus"
 9375..9637
 /note="AluSp repeat: matches 50..313 of consensus"

repeat_region 9638..9659
 /note="TIGER1 repeat: matches 2069..2090 of consensus"
 repeat_region 9696..11174
 /note="L1PA12 repeat: matches -1414..1133 of consensus"
 mRNA complement(11400..11692)
 /gene="dJ199L16.1"
 /note="match: cDNAs X59051 U66372 U14973 L31609 L31610;
 could be a pseudogene"
 /product="dJ199L16.1 (40S Ribosomal Protein S29 LIKE)"
 evidence=not_experimental
 complement(11400..11692)
 /gene="dJ199L16.1"
 /gene="dJ199L16.1"
 complement(11406..11411)
 /gene="dJ199L16.1"
 complement(11482..11634)
 /gene="dJ199L16.1"
 /note="match: proteins P14041 P54110 P30054 O15619 P26816
 O28368 O05635 P41057 P90983 O26125 P41058; could be a
 pseudogene"
 /codon_start=1
 /evidence=not_experimental
 /product="dJ199L16.1 (40S Ribosomal Protein S29 LIKE)"
 /protein_id="CAI8145.1"
 /db_xref="GI:3790138"
 /translation="FSQGSQSHSQVYSNMHGLIQKYLGNVCHQCFHGYLDIGFIKLDY
 SKRIIR"
 repeat_region 11801..11900
 /note="L1MC/D repeat: matches 5509..5619 of consensus"
 repeat_region 12006..12191
 /note="L1M4 repeat: matches 2657..3049 of consensus"
 repeat_region 12200..12599
 /note="MUT2B repeat: matches 1..400 of consensus"
 12566..12883
 /note="match: STS G08105"
 12595..12630
 /note="9 copies of tatg 100% conserved; differs from
 G08105"
 repeat_region 12606..12657
 /note="26 copies 2 mer at 7% conserved"
 12633..12676
 /note="11 copies of tcta 100% conserved; differs from
 G08105"
 repeat_region 12660..12726
 /note="MUT2B repeat: matches 384..448 of consensus"
 13405..13810
 /note="L2 repeat: matches 1259..1672 of consensus"
 14054..14262
 /note="MIR repeat: matches 20..216 of consensus"
 complement(14545..14834)
 /note="match: EST AA328142"
 14837..15005
 /note="MER20 repeat: matches 4..174 of consensus"
 15164..15275
 /note="L2 repeat: matches 2626..2750 of consensus"
 15342..15466
 /note="MIR repeat: matches 13..139 of consensus"
 15952..16120
 /note="MER5A repeat: matches 2..154 of consensus"
 16121..16317
 /note="MER45 repeat: matches 2..174 of consensus"
 16318..16351
 /note="MER5A repeat: matches 154..189 of consensus"
 16469..120086
 /note="L1PA2 repeat: matches 2519..6144 of consensus"
 20102..20214
 /note="L2 repeat: matches 1991..2102 of consensus"
 complement(22045..22272)
 /note="match: STS L42659"
 24543..24722
 /note="MUT1A1 repeat: matches 172..365 of consensus"
 24726..24874


```

repeat_region      /note="L1R15B repeat: matches 314. .464 of consensus"
25012. .25360
/note="L1R1A1 repeat: matches 19. .365 of consensus"
25361. .25643
repeat_region      /note="MER61C repeat: matches 134. .431 of consensus"
25691. .25878
repeat_region      /note="MER61B repeat: matches 1. .178 of consensus"
26111. .26322
repeat_region      /note="MIR repeat: matches 14. .260 of consensus"
27283. .27498
repeat_region      /note="L1P1A3 repeat: matches 5831. .6046 of consensus"
27499. .27555
repeat_region      /note="MER4D repeat: matches 410. .466 of consensus"
27659. .27810
repeat_region      /note="MER4D repeat: matches 522. .687 of consensus"
27847. .28049
repeat_region      /note="L1R29 repeat: matches 1. .202 of consensus"
28071. .28091
repeat_region      /note="L1M4A4 repeat: matches 6038. .6057 of consensus"
28351. .28644
repeat_region      /note="L1M5A repeat: matches 1. .293 of consensus"
28708. .28939
repeat_region      /note="L1P1A16 repeat: matches 5919. .6156 of consensus"
28963. .29076
repeat_region      /note="L1M4C4 repeat: matches 7857. .7975 of consensus"
29097. .29644
repeat_region      /note="L1M4C4 repeat: matches 7079. .7693 of consensus"
30797. .30977
repeat_region      /note="L1R8 repeat: matches 1. .182 of consensus"
30876. .31000
repeat_region      /note="MER87 repeat: matches 66. .190 of consensus"
31031. .31323
repeat_region      /note="L1R8 repeat: matches 397. .691 of consensus"
31418. .31451
repeat_region      /note="L17 copies 2 mer tt 82% conserved"
31454. .32540
repeat_region      /note="L1P2 repeat: matches 5058. .6144 of consensus"
32543. .32665
repeat_region      /note="L1M repeat: matches 5260. .5389 of consensus"
32879. .33080
repeat_region      /note="L1 repeat: matches 4085. .4292 of consensus"
33648. .33808
misc_feature        /note="match: STS AF020146"
34410. .34512
repeat_region      /note="L1P1A10 repeat: matches 6063. .6165 of consensus"
34639. .34702
repeat_region      /note="L32 copies 2 mer ga 70% conserved"
34839. .35073
repeat_region      /note="MIR repeat: matches 45. .256 of consensus"
34961. .35083
repeat_region      /note="L12 repeat: matches 2626. .2748 of consensus"
35350. .35423
repeat_region      /note="MIR repeat: matches 72. .142 of consensus"
37390. .37568

Query Match      2.6%: Score 38; DB 9; Length 134963;
Best Local Similarity 100.0%; Pred.No. 4.9e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GAGGTAAAAA...AAAAAAAAAAAAA 1463
DB 132649 GAGGTAAAAA...AAAAAAAAAAAAA 132612

```

ORGANISM

Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 145069)

REFERENCE

AUTHORS

Muzny,D.M., Adams,C., Adio-Odola,B., Alt-osman,F.R., Allen,C.,
Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbataia,J.,
Benton,D., Bimaga,K., Blankenburg,K., Bonini,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bulay,C.,
Burck,P., Burkett,C., Butrell,K.L., Byrd,N.C., Caron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Covle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Deigard,O.,
Dem,A.L., Ding,Y., Dinn,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgall,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hughes,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Joliver,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kralovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W.,
Louiaseg,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mashiney,E., Mcleod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,U., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peters,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojucokan,I., Rolfe,M.,
Riz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N.,
Slisson,I., Sodergren,E., Soneike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabot,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: HMOO
Center clone name: RP11-630D12
----- Summary Statistics -----
Sequencing vector: M13; I08821
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 148653 bases at least Q40
Consensus quality: 154105 bases at least Q30
Consensus quality: 158675 bases at least Q20
Estimated insert size: 153129; sum-of-contigs estimation
Estimated insert size: 208334; agarose-fp estimation
Quality coverage: 6.5x in Q20 bases; agarose-fp estimation
Quality coverage: 8.9x in Q20 bases; sum-of-contigs estimation

```
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
* 1 88843: contig of 88843 bp in length
* 88844 88943: gap of unknown length
* 88944 142825: contig of 53882 bp in length
* 142826 142925: gap of unknown length
* 142926 145069: contig of 2144 bp in length.
* Location/Qualifiers
* 1.145069
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="3"
* /clone="RP11-630D12"
BASE COUNT 48212 a 29567 c 26599 g 40487 t 204 others
ORIGIN
Query Match 2.6%; Score 38; DB 2; Length 145069;
Best Local Similarity 100.0%; Pred. No. 4,9e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 132109 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 132146

RESULT 24
AC092054/c 165894 bp DNA linear PRI 15-JAN-2002
LOCUS AC092054
DEFINITION Homo sapiens chromosome 3 clone RP11-437I21, complete sequence.
ACCESSION AC092054 AC069412
VERSION AC092054.2 GI:18151020
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 165894)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 165894)
Kaul, R.K., Olson, M.V., Raymond, C., Clendenning, J., Ivey, R.G. and
Haugen, E.D.
Direct Submission
Submitted (16-JUN-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 165894)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (15-JAN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jan 15, 2002 this sequence version replaced gi:14475987.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchtgs@u.washington.edu
Drafting Center: BCM
----- Project Information
Center Project name: chr-3
Center clone name: RP11-437I21 (bc0422)
----- Summary Statistics
Sequencing vector: M13; L08821; 38% of reads
```

Sequencing vector: plasmid: L08752; 62% of reads
Chemistry: Dye-Primer Bodipy; 12% of reads
Chemistry: Dye-terminator ET; 3% of reads
Chemistry: Dye-terminator Big Dye; 30% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165695 bases at least Q40
Consensus quality: 165862 bases at least Q30
Consensus quality: 165887 bases at least Q20
Insert size: 165894; sum-of-contigs
Quality coverage: 8.5x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': Mapping in progress
3': RP11-56P22 (UWGC:bc0176) AC093557

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
All regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.

HindIII		BglII		EcoRI	
SeqDerMap	FngPrInt	SeqDerMap	FngPrInt	SeqDerMap	FngPrInt
449	<800	6296	6426	8696	8758
6382	6500	2067	2138	6	<800
512	<800	4848	5072	3248	3212
449	<800	2640	2782	12449	12564
4200	4297	1241	1225	3524	3440
13	<800	4253	4360	1617	1613
14	<800	3592	3665	2055	2054
1996	2009	4422	4642	5062	5078
2216	2384	8187	8328	121	<800
3269	3295	2836	2993	2819	2931
10618	10502	2312	2496	1043	1018
892	861	3382	3502	8024	7950

619	<800	3279	3420	6680	6566	1424	1386	5453	5390	3767	3687
2743	2750	899	926	2785	2791	653	<800	10759	10721	4845	4848
2449	2497	656	<800	2400	2445	631	<800	1633	1642	2909	3054
4327	4297	3357	3502	1764	1791	2752	2750	2970	3097	359	<800
166	<800	507	<800	3064	3285	2681	2750	6554	6426	1186	1166
3226	3295	1994	2075	3900	3845	19	<800	3438	3579	853	872
1497	1475	527	<800	4935	4848	975	975			1871	1791
3208	3295	4216	4360	5344	5271	4157	4086				
2188	2164	1648	1642	999	1018						
239	<800	4284	4360	2222	2237						
586	<800	289	<800	1819	1791						
1874	1823	8248	8328	2741	2791						
652	<800	1406	1368	5859	5782						
274	<800	1861	1943	3444	3440						
772	861	11358	11428	1019	1018						
154	<800	448	<800	2749	2791						
1494	1475	1538	1502	2790	2791						
311	<800	297	<800	1211	1166						
1395	1386	1085	1077	3377	3440						
1156	1189	2526	2681	2493	2445						
170	<800	6144	6200	1159	1166						
1244	1189	3987	3730	373	<800						
239	<800	6763	6900	1490	1466						
4367	4297	5304	5584	2457	2445						
1208	1189	3197	3332	1154	1166						
2660	2750	2739	2864	1823	1791						
2142	2164	931	926	16	<800						
835	861	662	<800	6749	6758						
79	<800	4170	4360	443	<800						
5059	4966	4542	4642	387	<800						
10502	10502	1013	1019	4638	4604						
4292	4297	5739	5806	11487	11490						
1179	1189	169	<800	6881	6948						
1344	1386	401	<800	5635	5567						
1649	1619	43	<800	436	<800						
2141	2164	1482	1502	3315	3440						

Query Match 2.6%; Score 38; DB 9; Length 165894;
 Best Local Similarity 100.0%; Pred. No. 5e-10;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 Db 143412 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143375

RESULT 25
 AC027646 196560 bp DNA linear HTG 25-JAN-2002
 LOCUS
 DEFINITION Mus musculus clone RP23-185J14 strain C57BL6/J, WORKING DRAFT
 AC027646
 VERSION AC027646.9 GI:14149542
 KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 196560)
 REFERENCE Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Herrera,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.
 Direct Submission
 High Throughput Mouse Sequencing
 Unpublished
 2 (bases 1 to 196560)
 TITLE JOURNAL
 REFERENCE
 AUTHORS Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Herrera,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS Submitted (31-MAR-2000) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA
 On May 19, 2001 this sequence version replaced gi:14091719.
 COMMENT
 -----Genome Center
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: <http://wchanning.bwh.harvard.edu:9088/hpcg/jsp/hpcg/Sequence/mouse.html>
 Contact: gukmd@pepccod.bwh.harvard.edu
 -----Summary Statistics
 Center project name: ABE
 Sequencing vector: pUC18; L08752
 Chemistry: Dye-terminator Big Dye 3
 *Consensus quality: 190345 at least 020
 *Consensus quality: 189126 at least 030
 *Consensus quality: 187254 at least 040
 *Estimated insert size: agarose-FP - N/A
 *Estimated insert size: 196160 - sum-of-ctrls
 Quality coverage: agarose-FP - N/A
 Quality coverage: 8.6 x in Q20 bases; sum-of-ctrls estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

[illegible]

FEATURES

Location/Qualifiers

Source

1. 196560

/organism="Mus musculus"

```
/strain="C57BL6/J"
```

```
/db_xref="taxon:10090"
```

```
/clone="RP23-185J14"
```

```
/sex="male
```

1. .51804

```
/note="assembly_name:Contig221"
```

51825.90706

/note="assembly_name:Contig220

```
clone_end:T7
```

```
vector<side> left"
```

90727. 123879

```
/note="assembly name:Contiq219"
```

123900. 143293

```
/note="assembly name:Contig218"
```

143314 163801

```
170314: 100000
/notes="assembly name:Contia217"
```

7110LC addcmt+1
163822 173127

```
IOJ0ZZ: :1/31Z,  
(act&@="3ccomb]r namo:Cont+n216
```

71066 assembly_line

```
clone_end:3F0
```

vecloc_slue:11gml
173148 103603

1/3148. .182093

```
misc_feature      182714. 183958
                  /note="assembly_name:Contig214"
misc_feature      183979. 185019
                  /note="assembly_name:Contig213"
misc_feature      185040. 186528
                  /note="assembly_name:Contig212"
misc_feature      186549. 187616
                  /note="assembly_name:Contig211"
misc_feature      187637. 187921
                  /note="assembly_name:Contig210"
misc_feature      187942. 188538
                  /note="assembly_name:Contig209"
misc_feature      188559. 190007
                  /note="assembly_name:Contig208"
misc_feature      190028. 190585
                  /note="assembly_name:Contig207"
misc_feature      190606. 192178
                  /note="assembly_name:Contig206"
misc_feature      192199. 193302
                  /note="assembly_name:Contig205"
misc_feature      193323. 194502
                  /note="assembly_name:Contig204"
misc_feature      194523. 195116
                  /note="assembly_name:Contig203"
misc_feature      195137. 195681
                  /note="assembly_name:Contig202"
misc_feature      195702. 196560
```

	/note="assembly_name:contig201-			
BASE COUNT	57310	a	40079	c
ORIGIN	40312	g	58408	t
				451 others

Query Match	2.6%	Score 38;	DB 2;	Length 190500;
Best Local Similarity	100.0%	Pred. No. 5e-10;		
Matches 38;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1426	GAAGTAAAAAAAAAAAAAAAAAAAAAAAAA	1463
Db	44544	GAACTAAAAAAAAAAAAAAAAAAAAAAAAA	44581

RESULT 26					
AC020910	AC020910	203250 bp	DNA	linear	PRI 03-OCT-2001
LOCUS					
DEFINITION	Homo sapiens chromosome 19 clone CTD-2553L13, complete sequence.				
ACCESSION	AC020910				
VERSION	AC020910.6	GI:15887301			

KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
DOE Joint Genome Institute and Stanford Human Genome Center.
1 (bases 1 to 203201)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS

REFERENCE	2 (bases 1 to 203250)
JOURNAL	Unpublished
TITLE	Direct Submission
NUMBER	DOI: 10.1101/2020.07.20.201440

DOE JOINT GENOME INSTITUTE.
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE	3 (bases 1 to 203250)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (28-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell

REFERENCE
AUTHORS
TITLE
4 (bases 1 to 203250)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Drive, Walnut Creek, CA 94598, USA

JOURNAL Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Oct 3, 2001 this sequence version replaced gi:9558584.
COMMENT Location/Qualifiers

FEATURES
SOURCE

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2553L13"
1. .203201
/note="Draft Sequence Produced by DOE Joint Genome
Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sbgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence:
Estimated Total Number of Errors is 3.3.
SMS content:
SHGC-56626 G36949
SHGC-84051 G53736"

BASE COUNT 60596 a 42922 c 42178 g 57554 t
ORIGIN

Query Match 2.6%; Score 38; DB 9; Length 203250;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||
Db 86022 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 86059

RESULT 27
AC025528/c 238728 bp DNA linear HTG 25-JAN-2002
LOCUS
DEFINITION
DRAFT SEQUENCE, 54 unordered pieces.
AC025528
VERSION AC025528.9 GI:18092967
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 238728)
Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
Submitted (10-MAR-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
On Jan 9, 2002 this sequence version replaced gi:11094408.
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site:
http://wehanning.bwh.harvard.edu:9088/hpcg9/jsp/hpcg9/Sequence/mous
e.html
Contact: gntm@capcod.bwh.harvard.edu
-----Summary Statistics
Sequencing vector: pUC18; L88752
Chemistry: Dye-terminator Big Dye; 100%
*Consensus quality: 215188 at least Q20
*Consensus quality: 204865 at least Q30
*Consensus quality: 190041 at least Q40
Estimated insert size: agarose-FP - N/A
**Estimated insert size: 237668 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 3.6 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 54 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 31901: contig of 31901 bp in length
* 31902 31921: gap of unknown length
* 31922 63057: contig of 31136 bp in length
* 63058 63077: gap of unknown length
* 63078 83212: contig of 20135 bp in length
* 83213 83232: gap of unknown length
* 83233 96794: contig of 13562 bp in length
* 96795 96814: gap of unknown length
* 96815 111361: contig of 14547 bp in length
* 111362 120966: contig of 9585 bp in length
* 120967 120986: gap of unknown length
* 120987 128252: contig of 8266 bp in length
* 129253 129272: gap of unknown length
* 129273 136700: contig of 7428 bp in length
* 136701 136720: gap of unknown length
* 136721 143527: contig of 6807 bp in length
* 143528 143547: gap of unknown length
* 143548 151277: contig of 7730 bp in length
* 151278 151297: gap of unknown length
* 151298 157516: contig of 6219 bp in length
* 157517 157536: gap of unknown length
* 162065 162085: contig of 4529 bp in length
* 162086 162085: gap of unknown length
* 165464 165484: contig of 3379 bp in length
* 165485 165485: gap of unknown length
* 165486 167866: contig of 2382 bp in length
* 167867 170318: contig of 2432 bp in length
* 170319 170338: gap of unknown length
* 170339 174287: contig of 3949 bp in length
* 174288 174307: gap of unknown length
* 174308 175965: contig of 1662 bp in length
* 175966 175989: gap of unknown length
* 175990 180198: contig of 4209 bp in length
* 180199 180218: gap of unknown length
* 180219 181921: contig of 1703 bp in length
* 181922 181941: gap of unknown length
* 181942 184601: contig of 2660 bp in length
* 184602 184621: gap of unknown length
* 184622 187682: contig of 3061 bp in length
* 187683 187702: gap of unknown length
* 187703 190705: contig of 3003 bp in length
* 190706 190725: gap of unknown length
* 190726 193062: contig of 2337 bp in length
* 193063 193082: gap of unknown length
* 193083 195038: contig of 1956 bp in length
* 195039 195058: gap of unknown length
* 195059 196567: contig of 1509 bp in length
* 196568 196587: gap of unknown length
* 196588 197941: contig of 1354 bp in length
* 197942 197961: gap of unknown length
* 197962 198839: contig of 878 bp in length
* 198840 198859: gap of unknown length
* 198860 200490: contig of 1631 bp in length
* 200491 200510: gap of unknown length
* 200511 201707: contig of 1197 bp in length
* 201708 201727: gap of unknown length
* 201728 203928: contig of 2201 bp in length
* 203929 203948: gap of unknown length
* 203949 205346: contig of 1398 bp in length
* 205347 205366: gap of unknown length
* 205367 206975: contig of 1609 bp in length
* 206976 206995: gap of unknown length
* 206996 209493: contig of 2498 bp in length

```
* 209494 209513: gap of unknown length
* 209514 210401: contig of 888 bp in length
* 210402 210421: gap of unknown length
* 210422 212117: contig of 1696 bp in length
* 212118 212137: gap of unknown length
* 212138 213793: contig of 1656 bp in length
* 213794 213813: gap of unknown length
* 213814 216128: contig of 2315 bp in length
* 216129 216148: gap of unknown length
* 216149 217683: contig of 1535 bp in length
* 217684 217703: gap of unknown length
* 217704 218899: contig of 1196 bp in length
* 218900 218919: gap of unknown length
* 218920 220247: contig of 1328 bp in length
* 220248 220267: gap of unknown length
* 220268 220706: contig of 439 bp in length
* 220707 220726: gap of unknown length
* 220727 220973: contig of 247 bp in length
* 220974 220993: gap of unknown length
* 220994 222103: contig of 1110 bp in length
* 222104 222123: gap of unknown length
* 222124 222614: contig of 1491 bp in length
* 222615 223634: gap of unknown length
* 223635 225741: contig of 2107 bp in length
* 225742 225761: gap of unknown length
* 225762 227521: contig of 1760 bp in length
* 227522 227541: gap of unknown length
* 227542 228671: contig of 1130 bp in length
* 228672 228691: gap of unknown length
* 228692 230160: contig of 1469 bp in length
* 230161 230180: gap of unknown length
* 230181 231658: contig of 1478 bp in length
* 231659 231678: gap of unknown length
* 231679 233124: contig of 1446 bp in length
* 233125 233144: gap of unknown length
* 233145 234325: contig of 1181 bp in length
* 234326 234345: gap of unknown length
* 234346 236158: contig of 1813 bp in length
* 236159 237006: contig of 828 bp in length
* 237007 237026: gap of unknown length
* 237027 238728: contig of 1702 bp in length.

FEATURES
source
1. 238728
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-79P14"
/sex="male"
1. 31901
/note="assembly_name:Contig288"
misc_feature
31922..63057
/note="assembly_name:Contig287"
misc_feature
63078..83212
/note="assembly_name:Contig286"
83233..96794
/note="assembly_name:Contig285"
96815..111361
/note="assembly_name:Contig284"
111382..120966
/note="assembly_name:Contig283"
120987..129252
/note="assembly_name:Contig282"
129273..136700
/note="assembly_name:Contig281"
136721..143527
/note="assembly_name:Contig280"
143548..151277
/note="assembly_name:Contig279"
151298..157516
/note="assembly_name:Contig278"
157537..162065
misc_feature
```

```
/note="assembly_name:Contig277"
misc_feature
162086..165464
/note="assembly_name:Contig276"
clone_end:t7
vector_side:left"
misc_feature
165485..167866
/note="assembly_name:Contig275"
misc_feature
167887..170318
/note="assembly_name:Contig274"
misc_feature
170339..174287

Query Match 2.6%; Score 38; DB 2; Length 238728;
Best Local Similarity 100.0%; Pred. No. 5.1e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 19002 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 18965

RESULT 28
AX093307/c 396 bp DNA linear PAT 30-MAR-2001
LOCUS
DEFINITION Sequence 125 from Patent WO0118046.
ACCESSION AX093307
VERSION AX093307.1 GI:13509756
KEYWORDS
SOURCE
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 396)
AUTHORS Xu,J. and Stolk,J.A.
TITLE Ovarian tumor sequences and methods of use therefor
JOURNAL Patent: WO 0118046-A 125 15-MAR-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 396
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature
1. 396
/note="n = A,T,C or G"
BASE COUNT 115 a 71 c 60 g 125 t 25 others
ORIGIN
Query Match 2.5%; Score 37; DB 6; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 40 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4

RESULT 29
BC009480 445 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, SEC61, gamma subunit (S. cerevisiae), clone MGC:15510
DEFINITION IMAGE:3954682, mRNA, complete cds.
ACCESSION BC009480
VERSION BC009480.1 GI:14550464
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 445)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
```

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCID/CTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahy, Stephanie Ford, Julia Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 25 Row: 1 Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 4092055.

FEATURES
source
Location/Qualifiers
1..445
/organism="Homo sapiens"
/db_xref="LocustID:20335"
/db_xref="taxon:9606"
/clone="MGC:15310 IMAGE:3954682"
/tissue_type="skin, melanotic melanoma."
/clone_1lb="NIH_MGC_20"
/lab_host="DH10B-R"
/note="Vector: POTB7"
43..249
/codon_start=1
/product="SEC61, gamma subunit (S. cerevisiae)"
/protein_id="AAH0480.1"
/db_xref="GI:14550465"
/translation="MDQVMOFEPSPKQVKSIRLVKRCRTPDKRKEFGKIMATAIGF
ALMGIGFVKLIHIPINNIIYGC"
56 c 97 g 137 t

CDS
BASE COUNT 155 a 56 c 97 g 137 t
ORIGIN

Query Match 2.5%; Score 37; DB 9; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAA
Db 400 AAGTAAAAA
1463
436

RESULT 30
E01762
LOCUS E01762 508 bp RNA linear PAT 29-SEP-1997
DEFINITION Variant aequorin gene.
ACCESSION E01762
VERSION E01762.1 GI:2170015
KEYWORDS JP 1988291586-A/1.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 508)
AUTHORS Inoue,S., Kurose,K. and Sakaki,Y.
TITLE VARIANT Aequorin GENE
JOURNAL Patent: JP 1988291586-A 1 29-NOV-1988;
CHISSO CORP
OS Medusa
PN JP 1988291586-A/1
PD 29-NOV-1988
PF 23-MAY-1987 JP 1987126374
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC
C12N15/00//C07K13/00,C12P21/02,(C12P21/02,C12R1:19); CC
strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=PAQ440;
FH Key Location/Qualifiers

FH
FT misc.feature 1..958
FT FT
FT variation /note='aequorin gene'
FT replace(659,'c') /note='variant aequorin gene' FT variation
FT replace(590,'c') /note='variant aequorin gene' FT variation
FT replace(674,'c') /note='variant aequorin gene' FT variation
FT replace(569,'c') /note='variant aequorin gene' FT variation
FEATURES
source
Location/Qualifiers
1..508
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 191 a 75 c 133 g 109 t
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAA
Db 471 AAGTAAAAA
1463
507

RESULT 31
E01765
LOCUS E01765 508 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA sequence of equine.
ACCESSION E01765
VERSION E01765.1 GI:2170018
KEYWORDS JP 1988291593-A/1.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 508)
AUTHORS Inoue,S., Kurose,K. and Sakaki,Y.
TITLE PRODUCTION OF VARIANT Aequorin PROTEIN
JOURNAL Patent: JP 1988291593-A 1 29-NOV-1988;
CHISSO CORP
OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1988291593-A/1
PD 29-NOV-1988
PF 23-MAY-1987 JP 1987126373
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC
C12P21/02//C07K13/00,C12N15/00,(C12P21/02,C12R1:19); CC
strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH FT
FT CDS 1..958
FT FT /product='equine'.
FEATURES
source
Location/Qualifiers
1..508
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 191 a 75 c 133 g 109 t
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAA
Db 471 AAGTAAAAA
1463
507

```

RESULT 32
E08516          826 bp   RNA      linear      PAT 29-SEP-1997
DEFINITION      DNA encoding style-specific S-ribonuclease.
ACCESSION       E08516
VERSION          E08516.1 GI:2176631
KEYWORDS        JP 1994335389-A/1.
SOURCE          Lycopersicon peruvianum.
ORGANISM        Lycopersicon peruvianum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 826)
Tei,I., Nakada,K., Ito,T., Horiuchi,H., Ota,A., Takagi,M.,
Tsubura,H., Tanaka,H. and Ishiguro,T.,
S-RIBONUCLEASE SPECIFIC TO STYLE AND DNA SEQUENCE CODING THEREFOR
Patent: JP 1994335389-A 1 06-DEC-1994;
KAGOME CO LTD
OS      Lycopersicon peruvianum (tomato)
COMMENT  PN      JP 1994335389-A/1
          PD      06-DEC-1994
          PE      27-MAY-1993 JP 1993126286
          PI      TEI ITSUKIYON, NAKADA KENGO, ITO TORU, HORIUCHI HIROYUKI, PI
          OT      AKIMORI,
          PI      TAKAGI MASAMICHI, TSUBURA HIROKAZU, TANAKA HIROSHI, PI
          ISHIGURO YUKIO
          PC      C12N9/22,C12N15/52;
          CC      strandedness: Double;
          CC      topology: Linear;
          FH      Key      Location/Qualifiers
          FT      source      1..826
                          /organism='Lycopersicon peruvianum' FT
          FT      mat_peptide 1..486
                          /tissue_type='pistil'
          FT      /product='style-specific S-ribonuclease'.
FEATURES
source      1..826
            /organism='Lycopersicon peruvianum'
            /db_xref='taxon:4082'
BASE COUNT  309 a      150 c      131 g      236 t
ORIGIN
Query Match      2.5%; Score 37; DB 6; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      11111111111111111111111111111111
Db 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 33
E12586          826 bp   DNA      linear      PAT 24-JUN-1998
DEFINITION      cDNA encoding S-ribonuclease.
ACCESSION       E12586
VERSION          E12586.1 GI:3251418
KEYWORDS        JP 1997028381-A/3.
SOURCE          Lycopersicon peruvianum.
ORGANISM        Lycopersicon peruvianum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 826)
Tei,I., Minami,K. and Takagi,M.,
S-RIBONUCLEASE GENE AND PROMOTER SEQUENCE
Patent: JP 1997028381-A 3 04-FEB-1997;
TEI ITSUKIYON, MINAMI KOUICHI, TAKAGI MASAMICHI
OS      Lycopersicon peruvianum
COMMENT  PN      JP 1997028381-A/3

```

```

PD      04-FEB-1997
PE      24-JUL-1995 JP 1995187557
PI      TEI ITSUKIYON, MINAMI KOUICHI, TAKAGI MASAMICHI PC
C12N15/09,C07H21/04,C12N1/21//A01H1/00,C12N5/10,C12N9/22, PC
(C12N1/21,
PC C12N1.19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH      Key      Location/Qualifiers
FT      source      1..826
                          /organism='Lycopersicon peruvianum' FT   CDS
FT      1..489
                          /product='S-ribonuclease'
FT      3'UTR      /product='S-ribonuclease'
FT      Location/Qualifiers
FEATURES
source      1..826
            /organism='Lycopersicon peruvianum'
            /db_xref='taxon:4082'
BASE COUNT  309 a      150 c      131 g      236 t
ORIGIN
Query Match      2.5%; Score 37; DB 6; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      11111111111111111111111111111111
Db 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 34
S65048          826 bp   mRNA      linear      PLN 29-OCT-1993
LOCUS          self-incompatibility associated S-glycoprotein [clone II]
DEFINITION      [Lycopersicon peruvianum=tomatoes, PI 126441, style, Slla-plant,
ACCESSION       S65048
VERSION          S65048.1 GI:410002
KEYWORDS        Peruvian tomato style PI 126441 Slla-plant.
SOURCE          Lycopersicon peruvianum
ORGANISM        Lycopersicon peruvianum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 826)
Chung,I.K., Nakada,K., Tanaka,H., Ito,T., Horiuchi,H., Ohta,A. and
Takagi,M.,
Identification of cDNA clones coding for the style specific
Slla-glycoprotein gene associated with gametophytic
self-incompatibility in tomato (Lycopersicon peruvianum)
Biosci. Biotechnol. Biochem. 57 (7), 1172-1176 (1993)
93372490
GenBank staff at the National Library of Medicine created this
entry [NCBI gibsg 136760] from the original journal article.
This sequence comes from Fig. 1.
FEATURES
source      1..826
            /organism='Lycopersicon peruvianum'
            /db_xref='taxon:4082'
            /gene='self-incompatibility associated S-glycoprotein'
            /partial
            gene
BASE COUNT  309 a      152 c      129 g      236 t
ORIGIN
Query Match      2.5%; Score 37; DB 8; Length 826;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Qy 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||
 Db 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 35
 AF064030 829 bp mRNA linear PLN 22-MAR-1999
 LOCUS Helianthus tuberosus lectin 2 mRNA, complete cds.
 DEFINITION AF064030
 ACCESSION AF064030.1 GI:4091896
 VERSION
 KEYWORDS
 SOURCE Helianthus tuberosus.
 ORGANISM Helianthus tuberosus.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 1 (bases 1 to 829)
 Van Damme, E.J., Batte, A., Mazard, A.M., Verhaert, P., Horman, A.,
 Dedry, H., Rouge, P. and Peumans, W.J.
 Characterization and molecular cloning of the lectin from
 Helianthus tuberosus
 Eur. J. Biochem. 259 (1-2), 135-142 (1999)
 99115433
 2 (bases 1 to 829)
 Van Damme, E.J.M. and Peumans, W.J.
 Direct Submission
 Submitted (08-MAY-1998) Laboratory of Phytopathology and Plant
 Protection, Catholic University of Leuven, Willem de Croylaan 42,
 Leuven Herestraat, Brabant 3001, Belgium
 Location/Qualifiers
 1. 829
 /organism="Helianthus tuberosus"
 /db_xref="taxon:4233"
 82. 525
 /note="LECHEL2; agglutinin; helituba"
 /codon_start=1
 /product="lectin 2"
 /protein_id="AA011578.1"
 /db_xref="GI:4091897"
 /translation="MAASDTCVQAGPGCGGCKRMLOTAGCKTISITIKGCTIFSI
 QFVTKDNIETVLSGGQGVGQDKAETTFEADDDTYAISCTEAYVQMTVVSITFTOT
 NKVYGPFGVAGSRFSLPLTKGFAGFENSDVLDISIGVVVP"

BASE COUNT 258 a 150 c 188 g 233 t

ORIGIN

Query Match 2.5%; Score 37; DB 8; Length 829;
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||
 Db 785 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821

RESULT 36
 ATP27A 835 bp mRNA linear PLN 14-MAR-1997
 LOCUS A. thaliana mRNA for peroxidase ATP27a, partial.
 DEFINITION Y11792
 ACCESSION Y11792.1 GI:1890318
 VERSION
 KEYWORDS peroxidase.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 835)
 Weidner, K.G., Jespersen, H.M., Kjaersgaard, I.V.M., Justesen, A.F.,
 Oestergaard, L., Abelskov, A.K., Jensen, R.B., Hansen, L.N. and
 Rasmussen, S.K.
 From expressed sequence tags to structure, function, evolution and

JOURNAL
 REFERENCE 2 (bases 1 to 835)
 AUTHORS Justesen, A.F.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAR-1997) A.F. Justesen, University of Copenhagen,
 Dept. of Protein Chemistry, Oester Farimagsgade 2a, DK-1353
 Copenhagen K, DENMARK
 Location/Qualifiers
 1. 835
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="2011187"
 /cell_line="Columbia"
 /clone_1lb="lambda PR12"
 /dev_stage="roots"
 /dev_stage="stems"
 /dev_stage="rosettes"
 /dev_stage="flowers"
 /dev_stage="etiolated seedlings"
 <1. 616
 /EC_number="1.11.1.7"
 /codon_start=2
 /product="peroxidase ATP27a"
 /protein_id="CAA7248.1"
 /db_xref="GI:1890319"
 /db_xref="SPTREMBL:P93727"
 /translation="SCADVIAIAARDVYTLGGPWSVLGKRGDGTISANETRLNP
 PTEVNSLIQSFARGLSVKDMVTLDSGHTIGFSCHSSESLQNSKRFHDIDPSMNY
 AFAOTLKRCPTSNRGNAGTVLDSTSFVDVYKKQLSGVGSDQALLGDSRT
 KWIETPDAQOKAFREFEASVYKLNFGVKETGQVNTREVN"

BASE COUNT 285 a 172 c 166 g 212 t

ORIGIN

Query Match 2.5%; Score 37; DB 8; Length 835;
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||
 Db 768 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 804

RESULT 37
 BC018351 848 bp mRNA linear ROD 06-DEC-2001
 LOCUS Mus musculus
 DEFINITION Mus musculus (prosome, macropain) subunit, beta type 1,
 clone MGC:5916 IMAGE:3584538, mRNA, complete cds.
 ACCESSION BC018351
 VERSION BC018351.1 GI:17390819
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 848)
 Strausberg, R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/

REMARK
 COMMENT

Contact: villalonebcm.tmc.edu.
Villalón, D.R., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the T.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRK Plate: 10 Row: 1 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA g1: 7242196.

FEATURES

source

1. 848
/organism="Mus musculus"
/db_xref="LOCUSTID:19170"
/db_xref="taxon:10090"
/clone="MGC:5916 IMAGE:3584538"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_1ib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="vector: PCMV-SPORT6"
16. 738
/codon_start=1
/product="proteasome (prosome, macropain) subunit, beta
type 1"
/protein_id="AAH18351.1"
/db_xref="GI:17390820"
/translation="MSTVAAYRDERELQMGPHGAGVQLRFSPYAANGTVLAIAG
EDFSIVADPTRLSEGSFHTRDSPCYKLTDTVIGSGFGDCITLTKIIEARLKM
KHSNNAMTGTGIAAMLSTILTSRRFPYVYNIIGLDEEGKAVSFEDVGSYORD
SPKAGSASAMLOPLDNOVGFKNMQNEHVPLTLDRAMRLVKDFEISAARDVYTG
ALRICVTMEGRIREFVPLRKD"

CDS

BASE COUNT 259 a 179 c 214 g 196 t
ORIGIN

Query Match 2.5%; Score 37; DB 10; Length 848;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 795 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 831

RESULT 38
LOCUS AEUQA0440X 925 bp mRNA linear INV 18-MAR-1994
DEFINITION Aequorea victoria aequorin (AQ440) mRNA, complete cds.
ACCESSION 129571
VERSION 1.1 GI:461374
KEYWORDS aequorin; luminescent protein; oxygenase; photoprotein.
SOURCE Aequorea victoria cDNA to mRNA.
ORGANISM Aequorea victoria
Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
Aequoreidae; Aequorea.
1 (bases 1 to 925)
Inouye, S., Noguchi, M., Sakaki, Y., Takagi, Y., Miyata, T., Iwanaga, S.
and Tsuji, F. I
Cloning and sequence analysis of cDNA for the luminescent protein
aequorin
Proc. Natl. Acad. Sci. U.S.A. 82, 3154-3158 (1985)

TITLE
JOURNAL
MEDLINE
FEATURES
source
location/Qualifiers
1. 925
/organism="Aequorea victoria"
/db_xref="taxon:6100"
/cell_type="photocyte"
1. 925
/gene="AQ440"
82. 672
/gene="AQ440"
/note="EF-hand/Calcium binding site amino acids 31. 42
EF-hand/Calcium binding site amino acids 124. 133

EF-hand/Calcium binding site amino acids 160. 171"

/codon_start=1
/product="aequorin"
/protein_id="AA27720.1"
/db_xref="GI:461375"
/translation="MRSKQYSVKLSDFDPNPRIGRHKHMFNEDLVNHNKISLDEMY
YKASDIVINNIGATPEOAKRHRDAVEFGAGMYCVETDWPAYTEGKKLATPELE
KYAKNEPTLRIGWDALFDIKDONGATLDEMKAYTKAAGIIQSECEETFRVCD
IDSGOLDVDEMTROHLGFMYTMDPACERLYGAVP"

BASE COUNT 362 a 156 c 178 g 229 t
ORIGIN

Query Match 2.5%; Score 37; DB 3; Length 925;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 858 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 894

RESULT 39
LOCUS AR108851 958 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6113886.
ACCESSION AR108851
VERSION AR108851.1 GI:12825127
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 958)
AUTHORS Bryan, B.
TITLES Bioluminescent novelty items
JOURNAL Patent: US 6113886-A 5 05-SEP-2000;
FEATURES location/Qualifiers
1. 958
/organism="unknown"

BASE COUNT 363 a 156 c 210 g 229 t
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 40
LOCUS AR119320 958 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6152358.
ACCESSION AR119320
VERSION AR119320.1 GI:14102019
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 958)
AUTHORS Bryan, B.
TITLES Bioluminescent novelty items
JOURNAL Patent: US 6152358-A 5 28-NOV-2000;
FEATURES location/Qualifiers
1. 958
/organism="unknown"

BASE COUNT 363 a 156 c 210 g 229 t
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 41
AR151721 958 bp DNA linear PAT 08-AUG-2001
LOCUS
DEFINITION Sequence 5 from patent US 6232107.
ACCESSION AR151721
VERSION AR151721.1 GI:15117771
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 958)
Bryan,B.J. and Szent-Gyorgyi,C.
Luciferases, fluorescent proteins, nucleic acids encoding the
luciferases and fluorescent proteins and the use thereof in
diagnostics, high throughput screening and novelty items
Patent: US 6232107-A 5 15-MAY-2001;
Location/Qualifiers
1..958
/organism="unknown"
BASE COUNT 363 a 156 c 210 g 229 t
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 42
AX250563 958 bp DNA linear PAT 05-OCT-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0168824.
ACCESSION AX250563
VERSION AX250563.1 GI:15984303
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 958)
Bryan,B.J., Szent-Gyorgyi,C. and Szczepaniak,W.
Renilla reniformis fluorescent proteins, nucleic acids encoding the
fluorescent proteins and the use thereof in diagnostics, high
throughput screening and novelty items
Patent: WO 0168824-A 5 20-SEP-2001;
Protime, Ltd. (US); Bryan, Bruce J. (US)
Location/Qualifiers
1..958
/organism="Aequorea victoria"
/db_xref="taxon:6100"
115..705
/note="Apoecorin-in-encoding gene"
/codon_start=1
/protein_id="CAC93773.1"
/db_xref="GI:15984304"
/translation="MTSKQSYGVKLTSDFDNPRWIGRHKHMFELDVNHNKSLIDEMV
YKASDIYNNIGATPEQAKRRKRDVAFEGGAKMYGVETDMPATIEGKKLATDELE
YKAKNEPTLIRIMSDALFDIVKDDNGAITLDEWAAVYKAGIIOSSSDCEETFEVCD
IDESGQGLVDVMTKQHLGFWYTMDCERKLYGGAVP"
BASE COUNT 363 a 156 c 210 g 229 t
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 43
E01573 958 bp RNA linear PAT 29-SEP-1997
LOCUS
DEFINITION cDNA sequence of wild type of aquolin gene paQ440.
ACCESSION E01573
VERSION E01573.1 GI:2169826
KEYWORDS JP 1988098387-A/1.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 958)
Inoue,S., Kurose,K. and Sakaki,Y.
MUTANT AEQUORIN GENE
Patent: JP 1988098387-A 1 28-APR-1988;
CHISSO CORP
PN JP 1988098387-A/1
PD 28-APR-1988
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC
C12N15/00//C12P21/00,G01N21/76;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT misc.feature 1..958
/gene="wild type of aquolin gene paQ440".
Location/Qualifiers
1..958
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 363 a 156 c 210 g 229 t
ORIGIN

Query Match 2.5%; Score 37; DB 6; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 44
E01575 958 bp RNA linear PAT 29-SEP-1997
LOCUS
DEFINITION cDNA sequence of wild type of aquolin gene paQ440.
ACCESSION E01575
VERSION E01575.1 GI:2169828
KEYWORDS JP 1988098398-A/1.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 958)
Inoue,S., Kurose,K. and Sakaki,Y.
PRODUCTION OF MUTANT AEQUORIN PROTEIN
Patent: JP 1988098398-A 1 28-APR-1988;
CHISSO CORP
PN JP 1988098398-A/1
PD 28-APR-1988
PI INOUE SATOSHI, KUROSE KOICHI, SAKAKI YOSHIYUKI PC
C12P21/02,C12N15/00//C07K13/00,(C12P21/02,C12R1:19); CC

strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT misc_feature 1..958 /gene='wild type of aquolin gene PAQ440'.
Location/Qualifiers
source 1..958
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 363 a 156 c 210 g 229 t
ORIGIN

Query Match 2.58; Score 37; DB 6; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 45
E02319 958 bp RNA linear PAT 29-SEP-1997
LOCUS DNA encoding luminescent protein aqualine.
DEFINITION E02319
ACCESSION E02319.1 GI:2170555
VERSION JP 1990096597-A/1.
KEYWORDS Aequorea victoria.
SOURCE Aequorea victoria.
ORGANISM Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
Aequoreidae; Aequorea.
REFERENCE 1 (bases 1 to 958)
AUTHORS Takagi,Y., Sakaki,Y., Inoue,S., Noguchi,M., Iwanaga,S., Miyata,T.
and Furederitsuku,A.T.
TITLE BIOSYNTHESIS OF PHOTOPROTEIN AEQUORIN
JOURNAL Patent: JP 1990096597-A 1 09-Apr-1990;
CHISSO CORP
COMMENT OS Aequorea victoria
PN JP 1990096597-A/1
PD 09-APR-1990
PF 08-NOV-1988 JP 1989113000
PI TAKAGI YASUTAKA, SAKAKI YOSHIYUKI, INOUE SATOSHI, PI NOGUCHI MASATO,
PI IWANAGA SADAKI, MIYATA TOSHIYUKI, FUREDERITSUKU AI TSUJI PC
C07K13/00,C12P21/02//C12N15/12,(C12P21/02,C12R1:19); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=plaus;
CC *source: clone=paQ440;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FT 5'UTR 1..114
FT CDS 115..705
FT FT /product='luminescent protein aqualine' FT
FT mat_peptide 115..702
FT FT /product='luminescent protein aqualine' FT
FT 3'UTR 706..958
Location/Qualifiers
1..958
source Location/Qualifiers
1..958
/organism="Aequorea victoria"
/db_xref="taxon:6100"
BASE COUNT 362 a 155 c 212 g 229 t
ORIGIN

Query Match 2.58; Score 37; DB 6; Length 958;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

Search completed: November 5, 2002, 15:59:48
Job time : 4291.32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:52:52; Search time 174.666 Seconds
(without alignments)
14380.822 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463
Sequence: 1 caccagaatagctcgcgcgc.....aaaaaaaaaaaaaaaaaaaaa 1463

Scoring table:
OLIGO-NUC
Gapop 60.0, Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N.GeneSeq_032802.*

1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1463	100.0	1463	21	Maize Rad2/FEN-1 c
2	1360	93.0	1541	21	Maize Rad2/FEN-1 c
3	1267	86.6	1381	21	Maize Rad2/FEN-1 c
4	894	61.1	1478	21	Maize Rad2/FEN-1 c
5	40	2.7	584	21	Human prostate can
6	39	2.7	288	22	Human cervical can
7	39	2.7	305	22	Human prostate can
8	38	2.6	691	20	Human prostate tum
9	38	2.6	2381	21	Lung cancer associ

10	38	2.6	8711	22	AA526809	Human genomic DNA
11	37	2.5	302	21	AA598700	Human colon cancer
12	37	2.5	342	22	AA183214	Human polynucleoti
13	37	2.5	396	22	AA594934	Human ovarian can
14	37	2.5	411	22	AA188595	Human polynucleoti
15	37	2.5	412	22	AA183922	Human polynucleoti
16	37	2.5	426	22	AA19865	Human breast cance
17	37	2.5	524	21	AAA6571	Noncoding region o
18	37	2.5	699	22	AA564213	Human secreted pro
19	37	2.5	706	21	AA56563	Human secreted pro
20	37	2.5	826	16	AA079736	Flower style-speci
21	37	2.5	826	18	AA564553	Tomato S-ribonucle
22	37	2.5	957	11	AA004441	Aequorin gene, Ae
23	37	2.5	958	22	AA592233	Aequorin-encodi
24	37	2.5	958	22	AA592238	Aequorea victoria
25	37	2.5	959	9	AA591534	PAQ440 aequorin ge
26	37	2.5	1093	21	AA577825	Human cancer assoc
27	37	2.5	1159	22	AA502547	Human secreted pro
28	37	2.5	1252	22	AA59087	Human secreted pro
29	37	2.5	1293	24	AA563048	Cell death protect
30	37	2.5	1362	19	AA532421	Homo sapiens clone
31	37	2.5	1362	22	AA598437	Human cDNA clone C
32	37	2.5	1380	22	AA53111	Human secreted pro
33	37	2.5	1629	21	AA59580	Human secreted pro
34	37	2.5	1721	22	AA563820	Human secreted pro
35	37	2.5	1800	22	AA007579	Human secreted pro
36	37	2.5	2108	21	AA596570	A core 2 beta-1,6-
37	37	2.5	2198	21	AA53424	cDNA encoding huma
38	37	2.5	2260	22	AA197914	Human neuroblastom
39	37	2.5	2260	22	AA198068	Human neuroblastom
40	37	2.5	2297	21	AA523309	Sequence encoding
41	37	2.5	2602	21	AA12416	cDNA encoding a hu
42	37	2.5	2827	22	AA006812	Human LP8, a PDGF-
43	37	2.5	2836	19	AA540745	C. fells esterase,
44	37	2.5	2836	19	AA540745	C. fells esterase,
45	37	2.5	2836	19	AA540745	Ctenocephalides fe

ALIGNMENTS

RESULT 1	AAA27923	standard; cDNA: 1463 BP.
ID	AAA27923	
AC	AAA27923	
XX		
DT	12-SEP-2000	(first entry)
XX		
DE	Maize Rad2/FEN-1 cDNA.	
XX		
KW	Maize, Rad2/FEN-1; transgenic plant; male sterile plant;	
KW	endonuclease; exonuclease; DNA repair; gene targeting; ss.	
XX		
OS	Zea mays.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	85..1224
XX		/*tag= a
XX		
PN	WO200036109-A1.	
PD	22-JUN-2000.	
XX		
PF	16-NOV-1999;	99WO-US27147.
XX		
PR	15-DEC-1998;	98US-0112332.
XX		
PA	(PION-) PIONEER HI-BRED INT. INC.	
XX		
PI	Mahajan PB;	
XX		
DR	WPI; 2000-452026/39.	

P-PSDB: AAY95307.

Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting and the production of male sterile plants.

Claim 1; Page 69-71; 85bp; English.

The present sequence is that of maize cDNA coding for RAD2/FEN-1 (see AAY95307). The corresponding RNA was isolated from immature ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a structure specific endonuclease which under certain conditions also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to produce the Rad2/FEN-1 polypeptides in transgenic plant cells. The protein is involved in the regulation of DNA repair and recombination in plant systems and therefore may be used for improving gene targeting during further recombinant DNA protocols involving plants. RAD2/FEN-1 endonucleolytic activity is essential in DNA replication and nucleotide excision and repair reactions. The exolytic activity is involved in double strand break repair and joining. The protein is also useful in strand exchange reactions during homologous recombination. These functions may be useful in gene targeting and in the production of male sterile plants. The efficacy of gene targeting can be improved by the overexpression of exogenous Rad2/FEN-1 while male sterile plants can be produced by the down regulation of Rad2/FEN-1 expression.

Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other:

Query Match 100.0%; Score 1463; DB 21; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CACGAGATAGCTCGCGCGCGGTTCTTGGCCACTCCGGCTAGCCGCCCGCGCA 60
DB 1 CACGAGATAGCTCGCGCGCGGTTCTTGGCCACTCCGGCTAGCCGCCCGCGCA 60
QY 61 CCGGCGACGCGCGCGCGAGATGAGCATCAAGGTTTGACGAAATCTGCGCGAC 120
DB 61 CCGGCGACGCGCGCGCGAGATGAGCATCAAGGTTTGACGAAATCTGCGCGAC 120
QY 121 AATGCGCCCAAGGCGATGAGAGCAGAAAGTTGAGAGCTTCTGCGCGCAAAATCCG 180
DB 121 AATGCGCCCAAGGCGATGAGAGCAGAAAGTTGAGAGCTTCTGCGCGCAAAATCCG 180
QY 181 GTGAGCCCGAGCATGAGCATATACCATTCCTGATGTGTTGGAAGGACAGCATGGA 240
DB 181 GTGAGCCCGAGCATGAGCATATACCATTCCTGATGTGTTGGAAGGACAGCATGGA 240
QY 241 ACTCTCAAAATGAGCTGTGTAAGTCAGTACATTTGCAAGGAAATGTTCAACCGGACA 300
DB 241 ACTCTCAAAATGAGCTGTGTAAGTCAGTACATTTGCAAGGAAATGTTCAACCGGACA 300
QY 301 ATAGATTTACTGGAAGCGGGAATCAAGCCATTATGTTTGTATGCGCAAGCTCTGAT 360
DB 301 ATAGATTTACTGGAAGCGGGAATCAAGCCATTATGTTTGTATGCGCAAGCTCTGAT 360
QY 361 ATGAAGAAAGAAAGCTGTGTAAGATTAAGATCAAAAAAGATGATCAACCAAGATCTG 420
DB 361 ATGAAGAAAGAAAGCTGTGTAAGATTAAGATCAAAAAAGATGATCAACCAAGATCTG 420
QY 421 ACTGAGCAGTAGAGTAGAGATTAAGATGCGATTGAAAAATTTGAGCAAGAGACTGTA 480
DB 421 ACTGAGCAGTAGAGTAGAGATTAAGATGCGATTGAAAAATTTGAGCAAGAGACTGTA 480
QY 481 AAGGTCAAGAGCAACAAAGAAATTTGAAGCGCTATTAGACTTAAGGGGGTTCCCT 540
DB 481 AAGGTCAAGAGCAACAAAGAAATTTGAAGCGCTATTAGACTTAAGGGGGTTCCCT 540
QY 541 GTTGTAGAGGACCTTGTGAAGCAGAGCAAGATGTGCACCCCTTGCAATAAAGCATTAAG 600
DB 541 GTTGTAGAGGACCTTGTGAAGCAGAGCAAGATGTGCACCCCTTGCAATAAAGCATTAAG 600

```

```

QY 601 GNGTTCGCTGTTGCTTACAGAGATATGACATCCCTTACTTTTGGGGCTCCAGGTTCCCT 660
DB 601 GNGTTCGCTGTTGCTTACAGAGATATGACATCCCTTACTTTTGGGGCTCCAGGTTCCCT 660
QY 661 CGTCATTTAATGATCCAAAGTTCCAGAAAATACCTGTGATGGAATTTGATGTTCCCAAG 720
DB 661 CGTCATTTAATGATCCAAAGTTCCAGAAAATACCTGTGATGGAATTTGATGTTCCCAAG 720
QY 721 GTTTTGAGAGAGTTTAACCTCCATGACGAGCCAGTTTCATGATTTGTCATCTGTGTGA 780
DB 721 GTTTTGAGAGAGTTTAACCTCCATGACGAGCCAGTTTCATGATTTGTCATCTGTGTGA 780
QY 781 TGTGACTATTTGATGATGATCAATCAAGGATATCGGGGGCAACAGCTCTGAACTTATGCT 840
DB 781 TGTGACTATTTGATGATGATCAATCAAGGATATCGGGGGCAACAGCTCTGAACTTATGCT 840
QY 841 CAACATGGGCTTACCAAGAGAGCTGACGCTTGTCAAGAGCCTTAATGTCATTTGAT 900
DB 841 CAACATGGGCTTACCAAGAGAGCTGACGCTTGTCAAGAGCCTTAATGTCATTTGAT 900
QY 901 GAGGACTGGCTTACCAAGAGAGCTGACGCTTGTCAAGAGCCTTAATGTCATTTGAT 960
DB 901 GAGGACTGGCTTACCAAGAGAGCTGACGCTTGTCAAGAGCCTTAATGTCATTTGAT 960
QY 961 ATTCTGAGCTTAAATGAGACTGCACTGATGAGAGAGGCTCTCAATGTTTCTGTGTA 1020
DB 961 ATTCTGAGCTTAAATGAGACTGCACTGATGAGAGAGGCTCTCAATGTTTCTGTGTA 1020
QY 1021 GATTAATGTTTCAAGAGAGATGCGGTGACAAAGGCCATGAGAAATCAAAATTCGCAAG 1080
DB 1021 GATTAATGTTTCAAGAGAGATGCGGTGACAAAGGCCATGAGAAATCAAAATTCGCAAG 1080
QY 1081 AATTAATCTGTCGAAGAGAGACTGAGTCTTCTTTCAGGCAACTGCGCACATGACGA 1140
DB 1081 AATTAATCTGTCGAAGAGAGACTGAGTCTTCTTTCAGGCAACTGCGCACATGACGA 1140
QY 1141 CCGCTTAAACGGAAGAGAGACTTGGATTAACAAAGCAAGGACGCTGGAACAGAAACCA 1200
DB 1141 CCGCTTAAACGGAAGAGAGACTTGGATTAACAAAGCAAGGACGCTGGAACAGAAACCA 1200
QY 1201 AAGGCTGGTGAAGAGAAATATATCTTGGATGCTGATGATGACAACTAGCAAG 1260
DB 1201 AAGGCTGGTGAAGAGAAATATATCTTGGATGCTGATGATGACAACTAGCAAG 1260
QY 1261 CAGCGGTGCGGTGATCACTTCTGATTAATTAATTAACCTCTGTTTAACCTCAGAGCTTT 1320
DB 1261 CAGCGGTGCGGTGATCACTTCTGATTAATTAATTAACCTCTGTTTAACCTCAGAGCTTT 1320
QY 1321 GGTAAAGTTTGTCTCATGTTTCAAGCTGAGGATGATGTTGTTGAGAGATTTGCTG 1380
DB 1321 GGTAAAGTTTGTCTCATGTTTCAAGCTGAGGATGATGTTGTTGAGAGATTTGCTG 1380
QY 1381 TACCAAGTAAACAAACTATTCGCTGTTTACTTCTGTCCTTGAAGTAAAAA 1440
DB 1381 TACCAAGTAAACAAACTATTCGCTGTTTACTTCTGTCCTTGAAGTAAAAA 1440
QY 1441 AAAAAAAAAAAAAAAAAAAAAA 1463
DB 1441 AAAAAAAAAAAAAAAAAAAAAA 1463

```

RESULT 2

AAAZ7924 ID AAAZ7924 standard; cDNA: 1541 BP.

AAAZ7924:

12-SEP-2000 (first entry)

Maize Rad2/FEN-1 cDNA.

Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
endonuclease; exonuclease; DNA repair; gene targeting; ss.

```

XX OS Zea mays.
XX FH Key Location/Qualifiers
XX FT CDS 79..1218
XX FT /*tag= a
XX
XX PD WO200036109-A1.
XX PD 22-JUN-2000.
XX
XX PF 16-NOV-1999; 99MO-US27147.
XX PR 15-DEC-1998; 98US-0112332.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Mahajan PB;
XX DR WPI; 2000-452026/39.
XX DR P-PSDB; AAY95308.
XX
XX PT Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
XX PT recombination and repair in transgenic plants, e.g. for gene targeting
XX PT and the production of male sterile plants.
XX
XX PS Example 1; Page 73-74; 85pp; English.
XX
XX CC The present sequence is that of maize cDNA (ATCC PTA-533) coding
XX CC for RAD2/FEN-1 (see AAY95308). The corresponding RNA was isolated
XX CC from a B73 line seedling after a 10 day drought, heat shocked for
XX CC 10 hr, and allowed to recover under normal conditions. Rad2/FEN-1
XX CC is a structure specific endonuclease which under certain conditions
XX CC also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used
XX CC to produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
XX CC The protein is involved in the regulation of DNA repair and
XX CC recombination in plant systems and therefore may be used for
XX CC improving gene targeting during further recombinant DNA protocols
XX CC involving plants. Rad2/FEN-1 endonucleolytic activity is essential
XX CC in DNA replication and nucleotide excision and repair reactions.
XX CC The exolytic activity is involved in double strand break repair and
XX CC end joining. The protein is also useful in strand exchange
XX CC reactions during homologous recombination. These functions may be
XX CC useful in gene targeting and in the production of male sterile
XX CC plants. The efficacy of gene targeting can be improved by the
XX CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
XX CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX SQ Sequence 1541 BP; 473 A; 308 C; 377 G; 383 T; 0 other;

```

```

Query Match 93.0%; Score 1360; DB 21; Length 1541;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 GCGGTTCTTGGCCGCACTCGGCTCAGCGCCGCGCCGCAACCCGCGCAGACCCGCGAGA 80
DB 15 GCGGTTCTTGGCCGCACTCGGCTCAGCGCCGCGCCGCAACCCGCGCAGACCCGCGAGA 74
QY 81 CGAGATGGGCATCAAGGGTTTGACGAAATGCTGGCGAGACATGCGCCGAAGCGATGAA 140
DB 75 CGAGATGGGCATCAAGGGTTTGACGAAATGCTGGCGAGAACATGCGCCGAAGCGATGAA 134
QY 141 GAGGCGAAGATTGGAAGCTACTTGGCGCCGCAAAATGCGGTGAGCGCCAGCATGAGCAT 200
DB 135 GAGGCGAAGATTGGAAGCTACTTGGCGCCGCAAAATGCGGTGAGCGCCAGCATGAGCAT 194
QY 201 ATACCAATCTCTGATTTGATGTAAGGAGCAGCATGAAATCTCACAATGAAGGTGG 260
DB 195 ATACCAATCTCTGATTTGATGTAAGGAGCAGCATGAAATCTCACAATGAAGGTGG 254
QY 261 TGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGACAAATGAAGATTACTGGAAGCGG 320
DB 255 TGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGACAAATGAAGATTACTGGAAGCGG 314

```

```

QY 321 AATCAAGCAGTATTTATGTTTGTGATGGCAAGCCTCCTGATTTGAAGAACACAGACTTGC 380
DB 315 AATCAAGCAGTATTTATGTTTGTGATGGCAAGCCTCCTGATTTGAAGAACACAGACTTGC 374
QY 381 TAAAGATACCTCAAAAGAGATGATGCACCAAGATGTGACTGAGCAGATGAGGTGG 440
DB 375 TAAAGATACCTCAAAAGAGATGATGCACCAAGATGTGACTGAGCAGATGAGGTGG 434
QY 441 AGATTAAGATGCGATTGAAAAATTGAGCAGAGAGACTGTAAAGTCCACAGCAACACA 500
DB 435 AGATTAAGATGCGATTGAAAAATTGAGCAGAGAGACTGTAAAGTCCACAGCAACACA 494
QY 501 CGAAGATTGTAAAGCGCTATTAAGACTTATGAGGGTCTCTGTTAAGAGCAGCTTCTGA 560
DB 495 CGAAGATTGTAAAGCGCTATTAAGACTTATGAGGGTCTCTGTTAAGAGCAGCTTCTGA 554
QY 561 AGCAGAGCAGAAATGTGAGCCCTTTGCAATTAAGCATTAAGGTGCTGCTTGCATGAG 620
DB 555 AGCAGAGCAGAAATGTGAGCCCTTTGCAATTAAGCATTAAGGTGCTGCTTGCATGAG 614
QY 621 AGATATGAGACTCCCTTACTTTTGGGGCTCCAGGTTCCCTTGTGATTTAATGATCCAG 680
DB 615 AGATATGAGACTCCCTTACTTTTGGGGCTCCAGGTTCCCTTGTGATTTAATGATCCAG 674
QY 681 TTCCAGAAATATACCTGATGAGAAATTTGATGTTGCCAAGGTTTGGAGAGCTTGAAC 740
DB 675 TTCCAGAAATATACCTGATGAGAAATTTGATGTTGCCAAGGTTTGGAGAGCTTGAAC 734
QY 741 CACCATGAGACAGTTTCATTTGATTTGTCATCCTGTGATGATGATGATGATGATGAT 800
DB 735 CACCATGAGACAGTTTCATTTGATTTGTCATCCTGTGATGATGATGATGATGATGAT 794
QY 801 CAAAGTATGAGGGGGGCAAGCAGTCTGAAACTTATTCGTAACATGAGTGCATAGAAG 860
DB 795 CAAAGTATGAGGGGGGCAAGCAGTCTGAAACTTATTCGTAACATGAGTGCATAGAAG 854
QY 861 CATCTTGAGAAATTTATTAATGAAGACATATCAATTCCTGAGAGCTGGCCTTCCAGA 920
DB 855 CATCTTGAGAAATTTATTAATGAAGACATATCAATTCCTGAGAGCTGGCCTTCCAGA 914
QY 921 AGCTCGACGCTTTGTTCAAGAGACCTTAATGTCATTTGATTAATTCCTGATTAATGAC 980
DB 915 AGCTCGACGCTTTGTTCAAGAGACCTTAATGTCATTTGATTAATTCCTGATTAATGAC 974
QY 981 TGCACCTGATGAGGAGGCTCATTAAGTTCTGTTAAAGATATGCTTTCAACGAGAGA 1040
DB 975 TGCACCTGATGAGGAGGCTCATTAAGTTCTGTTAAAGATATGCTTTCAACGAGAGA 1034
QY 1041 TCGGGTGACAAAGGCCATAGAGAAGATCAAAATCGCCAAAGTAATTCGCAAGGAAG 1100
DB 1035 TCGGGTGAGAAAGGCCATAGAGAAGATCAAAATCGCCAAAGTAATTCGCAAGGAAG 1094
QY 1101 ACTTCGATCTTTTTCAGACCACTGCGACCAATTCGACACCGCTAAACGGAAGAGAC 1160
DB 1095 ACTTCGATCTTTTTCAGACCACTGCGACCAATTCGACACCGCTAAACGGAAGAGAC 1154
QY 1161 TTGCGATTAATAAACAAGCAGGCTGCGAACAAGAAACAAAGGCTGTGAAAAGAA 1220
DB 1155 TTGCGATTAATAAACAAGCAGGCTGCGAACAAGAAACAAAGGCTGTGAAAAGAA 1214
QY 1221 ATATCTTGGATGCTTGATGTACAACTACGACTACGAAACAGGCGGTGATCACTT 1280
DB 1215 ATATCTTGGATGCTTGATGTACAACTACGACTACGAAACAGGCGGTGATCACTT 1274
QY 1281 CGCTTATATTTAATTAATCCCTGTTTAACTCAGAGCTTTTGTTAAAGTTTGTCTAT 1340
DB 1275 CGCTTATATTTAATTAATCCCTGTTTAACTCAGAGCTTTTGTTAAAGTTTGTCTAT 1334
QY 1341 TCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATTGTTTGAAGAGATTGTTTGAAGAG 1400
DB 1335 TCAAGCTGGGGTAAAGTTAGTTGTTTGAAGAGATTGTTTGAAGAGATTGTTTGAAGAG 1394

```


OY	1204	GCCTGGTGGAAAGGAAGAATAATCCTGGACGTTGSAATGTACAACACTCGCACTACGAAACGAG	1263
OY	1204	GCCTGGTGGAAAGGAAGAATAATCCTGGACGTTGSAATGTACAACACTCGCACTACGAAACGAG	1264
Dd	1156	GCTGTGTGGAAGAAGAAATATCATCTTGTCGATGCTTGAATGTACACTACGACTACGAAAGCAG	1215
OY	1264	CGGTGGCGGTGATCACCCTCGCTTAGATTATTTAATCTCCCTGTTTTTAACTCAGAGCTTTGGT	1323
OY	1264	CGGTGGCGGTGATCACCCTCGCTTAGATTATTTAATCTCCCTGTTTTTAACTCAGAGCTTTGGT	1324
Dd	1216	CGGTGGCGGTGATCACCCTCGCTTAGATTATTTAATCTCCCTGTTTTTAACTCAGAGCTTTGGT	1275
OY	1324	AAAAGTT 1330	
Dd	1276	AAAAAGTT 1282	
		RESULT 4	
		AAA27926	
		AAA27926 standard; cdna; 1478 BP.	
XX	AAAZ7926;		
XX	AC		
XX	DT	12-SEP-2000 (first entry)	
XX	DE	Maize Rad2/FEN-1 cDNA.	
XX	KW	Maize; Rad2/FEN-1; transgenic plant; male sterile plant;	
XX	RW	endonuclease; exonuclease; DNA repair; gene targeting; ss.	
XX	Zea mays.		
XX	CS		
FH	Key	Location/Qualifiers	
FT	CDS	97..1236	
FT		/tag= a	
XX			
PN	MO200036109-A1.		
PD	22-JUN-2000.		
XX			
PF	16-NOV-1999; 99MO-US27147.		
XX			
PR	15-DEC-1998; 98US-0112332.		
XX			
PA	(PION-) PIONEER HI-BRED INT INC.		
XX			
PJ	Mahajan PB;		
DR	MP1: 2000-452026/39.		
DR	P-P5DB; AAY95310.		
PT			
PT	Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA		
PT	recombination and repair in transgenic plants, e.g. for gene targeting		
XX	and the production of male sterile plants -		
PS			
PS	Example 1; Page 77-79; 85pp; English.		
XX			
XX	The present sequence is that of maize cDNA coding for RAD2/FEN-1		
CC	(see AAY95310). The cDNA was isolated from a library prepared		
CC	from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a		
CC	structure specific endonuclease which under certain conditions also		
CC	acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to		
CC	produce the Rad2/FEN-1 polypeptides in transgenic plant cells.		
CC	The protein is involved in the regulation of DNA repair and		
CC	recombination in plant systems and therefore may be used for		
CC	improving gene targeting during further recombinant DNA protocols		
CC	involving plants. RAD2/FEN-1 endonucleolytic activity is essential		
CC	in DNA replication and nucleotide excision and repair reactions.		
CC	The exoinit activity is involved in double strand break repair and		
CC	end joining. The protein is also useful in strand exchange		
CC	reactions during homologous recombination. These functions may be		
CC	useful in gene targeting and in the production of male sterile		
CC	plants. The efficacy of gene targeting can be improved by the		
CC	overexpression of exogenous Rad2/FEN-1 while male sterile plants can		
CC	be produced by the down regulation of Rad2/FEN-1 expression.		
SO	Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other:		
XX			

Query Match	61.1%;	Score 894;	DB 21;	Length 1478;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1244;	Conservative	0;	Mismatches 7;	Indels 0;
				Gaps 0;

QY	20	CGGGTTTTCTGGGGCCACTCCGGGCTACGGCGCGCGCCGGACCCGGCCACAGCCGCCAG	79
Db	32	CGGGTTTTCTGGGGCCACTCCGGGCTACGGCGCGCGCCGGACCCGGCCACAGCCGCCAG	91
QY	80	ACGAGATGGGCATCAAGGTTTGTACCAAACTGCTGGCGGACAMTGCGCCCAAGCGATGA	139
Db	92	ACGAGATGGGCATCAAGGTTTGTACCAAACTGCTGGCGGACAMTGCGCCCAAGCGATGA	151
QY	140	AGGAGCAGAATTTCGAGAGCTACTTTGGCGCCGCAAAATGCGCGTGCAGGCCAGCATGAGCA	199
Db	152	AGGAGCAGAATTTCGAGAGCTACTTTGGCGCGCAAAATGCGCGTGCAGGCCAGCATGAGCA	211
QY	200	TATACCAAGTTCCTGATTGTAGTTTGGAAGGACAAGCTGTGAAACTCTCAACAAATGAAGCTG	259
Db	212	TCTACCAAGTTCCTGATTGTAGTTTGGAAGGACAAGCTGTGAAACTCTCAACAAATGAAGCTG	271
QY	260	GTGAAGTCACATGTCATTTTGCAGAGAAATGTTCAACCGGCAACTTAAGATTACTGCAAGCGG	319
Db	272	GTGAAGTCACATGTCATTTTGCAGAGAAATGTTCAACCGGCAACTTAAGATTACTGCAAGCGG	331
QY	320	GAATCAAGCCAGTTTATGTTTTGATGGCAAGCCCTCTATATGTAAGAACAAGAGCTTG	379
Db	332	GAATCAAGCCAGTTTATGTTTTGATGGCAAGCCCTCTATATGTAAGAACAAGAGCTTG	391
QY	380	CTAAAGATACTCAAAAAAGATGATGCAACCAAAATCTGACTGAGGCAGTAAGAGTGA	439
Db	392	CTAAAGATACTCAAAAAAGATGATGCAACCAAAATCTGACTGAGGCAGTAAGAGTGA	451
QY	440	GAGATAAAGATGCATGGAATAAATTGAGCAGAAGACTCTAAAGGTCACAAGCAACACA	499
Db	452	GAGATAAAGATGCATGGAATAAATTGAGCAGAAGACTCTAAAGGTCACAAGCAACACA	511
QY	500	ACGAAGATTGTAAACGGCTTTTAAAGACTTTTGGGGGTCCTGTGTAAAGAGCACTTCGTG	559
Db	512	ACGAAGATTGTAAACGGCTTTTAAAGACTTTTGGGGGTCCTGTGTAAAGAGCACTTCGTG	571
QY	560	AAGCAGAAGCAGAAATGTGCGAGCCCTTGTGCATTAACGATTAAGGTGTGTGCTTGCATG	619
Db	572	AAGCAGAAGCAGAAATGTGCGAGCCCTTGTGCATTAACGATTAAGGTGTGTGCTTGCATG	631
QY	620	AAGATATGGACTCCCTTACTTTTGGGGGTCACAGGTTCTTGTGATTTAATGATGCCAA	679
Db	632	AAGATATGGACTCCCTTACTTTTGGGGGTCACAGGTTCTTGTGATTTAATGATGCCAA	691
QY	680	GTTGCCAAGAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTTGAGAGAGCTTGAAAC	739
Db	692	GTTGCCAAGAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTTGAGAGAGCTTGAAAC	751
QY	740	TCACCATGAGACCACTGATTTGATTTGTGCACTCTGTGTGATGTGACTATTGTGTACGA	799
Db	752	TCACCATGAGACCACTGATTTGATTTGTGCACTCTGTGTGATGTGACTATTGTGTACGA	811
QY	800	TCAAAGGATATGGGGGGCAAAAGACTGTCAAACTTATTTGTCACATGAGGTCATATGAAA	859
Db	812	TCAAAGGATATGGGGGGCAAAAGACTGTCAAACTTATTTGTCACATGAGGTCATATGAAA	871
QY	860	GCATCTTGGAGCATCTTAATTAAGACAGATATCAAAATTCCTGAGCACTGGCTTACCAAG	919
Db	872	GCATCTTGGAGCATCTTAATTAAGACAGATATCAAAATTCCTGAGCACTGGCTTACCAAG	931
QY	920	AAGTCGACGCTTTTCAAGAGAGCTATGTCACATGATATTCCTGAGCTAAATATGCA	979
Db	932	AAGTCGACGCTTTTCAAGAGAGCTATGTCACATGATATTCCTGAGCTAAATATGCA	991
QY	980	CTGCACTGATGAGAGGGTCTATAAGTTTCTGCGTAAAGATATGGTTTCAACGAG	1039
Db	992	CTGCACTGATGAGAGGGTCTATAAGTTTCTGCGTAAAGATATGGTTTCAACGAG	1051

QY 1040 ATCGGGTACAAAGGCCATAGAGATCAATCTGCCAAAGATTAATGCTGCAAGGAA 1099
|||||
DB 1052 ATCGGGTACAAAGGCCATAGAGATCAATCTGCCAAAGATTAATGCTGCAAGGAA 1111
QY 1100 GACTCGAGTCCCTTTTTCACAGCACTGCCACCATCAGCACCGCTAAACGGAAGAGA 1159
|||||
DB 1112 GACTCGAGTCCCTTTTTCACAGCACTGCCACCATCAGCACCGCTAAACGGAAGAGA 1171
QY 1160 CTTCGGATTAACACAGCAGCGCTGCCAAGCAAAAGCAAGCTGCTGGAAGAGCA 1219
|||||
DB 1172 CTTCGGATTAACACAGCAGCGCTGCCAAGCAAAAGCAAGCTGCTGGAAGAGCA 1231
QY 1220 AATAATCTTGATGCTTGATGTACAACTAGACTACGAAAGCAGCGGTGC 1270
|||||
DB 1232 AATAATCTTGATGCTTGATGTACAACTAGACTACGAAAGCAGCGGTGC 1282

RESULT 5

AAFI5908
ID AAFI5908 standard; cDNA; 584 BP.

AC AAFI5908;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:343.

XX Human: prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytosolic; cardiovascular; immunomodulatory; muscular;
XX vulnary; gastrointestinal; nephrologic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease; ss.

OS Homo sapiens.

XX WO20005174-A1.

PN 21-SEP-2000.

PD 08-MAR-2000; 2000MO-US05988.

PE 12-MAR-1999; 99US-0124270.

PR (HUMA-) HUMAN GENOME SCI INC.
(ROSE/) ROSEN C A.

PA Rosen CA, Ruben SM;

PI WPI; 2000-587513/55.

PS P-PSDB; AAB56705.

PT Prostate cancer associated gene sequences, referred to as prostate
XX cancer antigens, useful for treatment, prevention, and diagnosis of
XX disorders such as prostate cancer -

XX Claim 1; Page 846; 2338bp; English.

XX AAFI566 to AAFI6505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytosolic,
XX cardiactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX nephrologic, antiinfective, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention.

SO Sequence 584 BP; 188 A; 120 C; 110 G; 165 T; 1 other;

Query Match 2.7%; Score 40; DB 21; Length 584;

Best Local Similarity 100.0%; Pred. No. 4.7e-06;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 TTGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1463
|||||

DB 543 TTGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 582

RESULT 6

AAH69821/C
ID AAH69821 standard; cDNA; 288 BP.

AC AAH69821;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 1095.

XX Cervical cancer; cytosolic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX WO200142467-A2.

PN 14-JUN-2001.

PD 08-DEC-2000; 2000MO-US33312.

PE 08-DEC-1999; 99US-0166681.

PR 21-DEC-1999; 99US-0171350.

PR 14-MAR-2000; 2000US-0189315.

PR 12-MAY-2000; 2000US-0203791.

PR 09-JUN-2000; 2000US-0210600.

PR 21-JUL-2000; 2000US-0220114.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PA Schlegel R, Deeds J, Berger A, Zhao X;

PI WPI; 2001-375006/39.

PS Claim 1; Page 279; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytosolic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy.

XX Sequence 288 BP; 88 A; 49 C; 42 G; 107 T; 2 other;

SO Query Match 2.7%; Score 39; DB 22; Length 288;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1423 TTGGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1461
|||||

DB 66 TTGGAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 28

RESULT 7

AAH71246/C
ID AAH71246 standard; cDNA; 305 BP.

AC AAH71246;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 1095.

XX Cervical cancer; cytosolic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX WO200142467-A2.

AC	AAH71246;	
XX		
DT	19-SEP-2001	(first entry)
XX		
DE	Human cervical cancer marker nucleic acid 2520.	
XX		
KW	Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.	
XX		
OS	Homo sapiens.	
PN	WO200142467-A2.	
PD	14-JUN-2001.	
XX		
PF	08-DEC-2000; 2000WO-US33312.	
XX		
PR	08-DEC-1999; 99US-0169681.	
PR	21-DEC-1999; 99US-0171350.	
PR	14-MAR-2000; 2000US-0189315.	
PR	12-MAY-2000; 2000US-0203791.	
PR	09-JUN-2000; 2000US-0210600.	
XX	21-JUL-2000; 2000US-0220114.	
PA	(MILL-) MILENNIUM PREDICTIVE MEDICINE INC.	
PI	Schlegel R, Deeds J, Berger A, Zhao X;	
XX		
DR	WPI; 2001-375006/39.	
XX		
PT	New isolated nucleic acid for diagnosing and treating cervical cancer	
XX	and for assessing and detecting compounds for treating the cancer -	
XX		
PS	Claim 1: Page 525; 1051pp; English.	
XX		
CC	The invention relates to novel genes (AAH68727-AAH73383) associated with	
CC	cervical cancer with cyrostatic activity. The nucleic acids and encoded	
CC	polypeptides are useful: to assess if a patient is afflicted with	
CC	cervical cancer or has a pre-malignant condition; to monitor the	
CC	progression of cervical cancer or a premalignant condition in a patient;	
CC	and to select and/or assess the efficacy of a compound or therapy for	
CC	inhibiting cervical cancer in a patient. The nucleic acids may also be	
CC	useful for gene therapy.	
XX		
SO	Sequence 305 BP; 96 A; 45 C; 51 G; 113 T; 0 other;	
XX		
Query Match	2.7%; Score 39; DB 22; Length 305;	
Best Local Similarity	100.0%; Pred. No. 1.3e-05;	
Matches 39; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
QY	1423 TTTGAAGTAAAAA	1461
DB	44 TTGCACTAAAAA	6
RESULT 8		
AAZ52929		
ID	AAZ52929 standard; CDNA: 691 BP.	
XX		
AC	AAZ52929;	
XX		
DT	14-MAR-2000 (first entry)	
XX		
DE	Human prostate tumor cDNA library derived EST fragment #72.	
XX		
KW	Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;	
KW	treatment; ds.	
OS	Homo sapiens.	
XX		
PN	DE19820190-A1.	
XX		
PD	04-NOV-1999.	
XX		

```

PF 28-APR-1998: 98DE-1020190.
xx
PR 28-APR-1998: 98DE-1020190.
xx
PA (META-) METAGEN GES GENOMFORSCHUNG MHH.
xx
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
xx
DR WPI: 1999-621386/54.
xx
DR P-PSDB: AAY74028, AAY74029, AAY74030.
xx
PT New human nucleic acid sequences from pancreatic tumors, and related
xx proteins -
xx
PS Claim 2: Page 240-241; 502pp; German.
xx
CC This invention describes novel polypeptides and their encoding nucleic
xx acids derived from human pancreatic tumor tissue which have cytostatic
xx activity. The sequences are also useful in producing pharmaceutical
xx compositions for treatment of pancreatic tumors. AA252858-253014
xx represent expressed sequence tag (EST) fragments derived from a human
xx pancreatic tumor cDNA library and which encode the proteins represented
xx in AAY73814-Y74252.
xx
SQ Sequence 691 BP; 192 A; 181 C; 174 G; 144 T; 0 other;
xx
Query Match 2.6%; Score 38; DB 20; Length 691;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Db 631 GAACGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 668

RESULT 9
AAFI8106
ID AFI8106 standard; DNA: 2381 BP.
xx
AC AFI8106;
xx
DT 14-MAR-2001 (first entry)
xx
DE Lung cancer associated polynucleotide sequence SEQ ID 125.
xx
KW Human; lung cancer associated protein; neuroprotective; cytostatic;
xx cardiocactive; immunomodulatory; muscular active; vulnerrary;
xx gastrointestinal; nephrotoxic; antinfctive; gynecological;
xx antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
xx proliferative disorder; wound healing; infectious disease; ds.
xx
OS Homo sapiens.
xx
PN W0200055180-A2.
xx
PD 21-SEP-2000.
xx
PE 08-MAR-2000; 2000WO-0505918.
xx
PR 12-MAR-1999; 99US-0124270.
xx
PA (HUMA-) HUMAN GENOME SCI INC.
xx PA (ROSE/) ROSEN C A.
xx
PI Ruben SM;
xx
DR WPI: 2000-587514/55.
xx
DR P-PSDB: AAB58230.
xx
PT Lung cancer associated gene sequences, referred to as lung cancer
xx antigens, useful for treatment, prevention, and diagnosis of disorders
xx such as lung cancer -
xx

```

PS Claim 1: Page 593; 1425bp; English.
XX
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardiactive; gastro
CC immunomodulatory; muscular active general; vulnary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.
XX
SQ Sequence 2381 BP; 527 A; 704 C; 659 G; 486 T; 5 other;
XX
Query Match 2.6%; Score 38; DB 21; Length 2381;
Best Local Similarity 100.0%; Pred.No. 2.1e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1426 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2272 GAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2309
RESULT 10
AAS26809
ID AAS26809 standard; DNA; 8711 BP.
XX
AC AAS26809;
XX
DT 07-NOV-2001 (first entry)
XX
XX Human genomic DNA encoding partial novel secreted protein, Seq ID 1783.
DE
XX Human: immunosuppressive; antiarthritic; ds; antipneumatic;
KW cytosolic; cardiant; vasotropic; cerebroprotective; nocrotic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX WO200155322-A2.
XX
PD 02-AUG-2001.
PF 17-JAN-2001; 2001MO-US01341.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0186874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216860.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225457.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239355.
PR 13-OCT-2000; 2000US-0239397.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

CC	hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC	cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC	e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
CC	Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC	and ocular disorders e.g. corneal infection, and many other
CC	disorders listed in the specification. The polypeptides can also
CC	be used to aid wound healing and epithelial cell proliferation, to
CC	prevent skin aging due to sunburn, to maintain organs before
CC	transplantation, for supporting cell culture of primary tissues, to
CC	regenerate tissues and in chemotaxis. The polypeptides can also be used
CC	as a food additive or preservative to increase or decrease storage
CC	capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC	minerals, cofactors and other nutritional components. The present
CC	sequence is a genomic DNA encoding a partial novel secreted protein of
CC	the invention.
Query Match:	2.6%; Score 38; DB 22; Length 8711;
Best Local Similarity	100.0%; Pred. No. 1.7e-05;
Matches 39; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1426 GAAGTAAAAAAAAAAAAAAAAAAAAA 1463 Db 6402 GAAGTAAAAAAAAAAAAAAAAAAAAA 6519
RESULT 11	
ID	AAC98700 standard; cDNA; 302 BP.
XX	AAC98700:
XX	09-MAR-2001 (First entry)
DE	Human colon cancer antigen nucleotide sequence SEQ ID NO:710.
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection; identification; cytostatic; cardioactive; neuroprotective; vulnery; immunomodulatory; muscular; gynaecological; gastrointestinal; neurotropic; antiinfective; antibacterial; gene therapy; wound; neural disorder; immune system disorder; muscular disorder; reproductive disorder; gastrointestinal disorder; renal disorder; infectious disease; cardiovascular disorder; ss.
OS	Homo sapiens.
PN	MO200055351-A1.
XX	21-SEP-2000.
XX	08-MAR-2000; 2000WO-US05883.
XX	12-MAR-1999; 99US-0124270.
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Rosen CA, Ruben SM;
DR	WPI: 2000-587534/55. P-PDB: AAB53943.
PT	Colon cancer associated gene sequences, referred to as colon cancer antigens, useful for the treatment, prevention, and diagnosis of colon disorders such as colon cancer -
PS	Claim 1; Page 1268; 2104pp; English.
CC	AAC97991 to AAC98763 encode the human colon cancer associated proteins, called human colon cancer antigens, given in AAB5334 to AAB54006. The human colon cancer antigens can have cytostatic, cardioactive, muscular; neuroprotective, immunomodulatory, gynaecological, gastrointestinal, vulnerable, neurotropic, antiinfective and antibacterial activities, and can be used in gene therapy. The colon cancer antigen polynucleotides, proteins and antibodies

CC treatment and diagnosis of colon disorders, such as colon cancer. The
 CC polynucleotides may be used in diagnostics and research, such as for
 CC chromosome identification, and as hybridisation probes. The proteins
 CC may also be used to prevent diseases such as neural disorders, immune
 CC system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, wounds, renal disorders, infectious
 CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
 CC AAB54007 represent sequences used in the exemplification of the present
 CC invention.

CC Sequence 302 BP; 159 A; 20 C; 35 G; 86 T; 2 other;

XX SQ

Query Match 2.5%; Score 37; DB 21; Length 302;
 Best Local Similarity 100.0%; Pred. NO. 7.4e-05;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||
 DB 207 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 243

RESULT 12
 AAI83214
 ID AAI83214 standard; cDNA: 342 BP.

XX AAI83214;
 AC
 XX 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 3274.
 XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-514838/56.
 DR P-PDB; AAO03283.
 XX

PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX

PS Claim 1; SEQ ID NO 3274; 1399pp + Sequence Listing: English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulatory
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 342 BP; 149 A; 52 C; 74 G; 67 T; 0 other;
 SQ

Query Match 2.5%; Score 37; DB 22; Length 342;
 Best Local Similarity 100.0%; Pred. NO. 7.2e-05;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||
 DB 54 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 90

RESULT 13
 AAF94934/C
 ID AAF94934 standard; cDNA: 396 BP.

XX AAF94934;
 AC
 XX 23-MAY-2001 (first entry)
 DE Human ovarian cancer associated coding sequence SEQ ID NO: 125.
 DE
 XX Human; ovarian cancer; vaccine; gene therapy; carcinoma; ss.
 KW
 KW Human; ovarian cancer; vaccine; gene therapy; carcinoma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200118046-A2.
 PD 15-MAR-2001.
 XX
 PF 08-SEP-2000; 2000WO-US24827.
 XX
 PR 10-SEP-1999; 99US-0394374.
 PR 01-MAY-2000; 2000US-0561778.
 PR 15-AUG-2000; 2000US-0640173.
 PR 07-SEP-2000; 2000US-0656668.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Stolk JA;
 XX WPI; 2001-211395/21.
 DR
 XX
 PF Isolated polypeptides associated with ovarian carcinomas, and the
 PT nucleic acids that encode them, useful for the prevention diagnosis and
 PT treatment of ovarian cancers -
 XX

PS Claim 5; Page 157; 189pp; English.
 XX
 CC The present invention provides a number of coding sequences and proteins,
 CC the over-expression of which is associated with ovarian carcinoma/cancer.
 CC These can be used in the diagnosis, treatment and prevention of ovarian
 CC cancer, optionally by gene therapy or in the form of a vaccine. The
 CC present sequence is an example of one of these sequences.
 XX

SQ Sequence 396 BP; 115 A; 71 C; 60 G; 125 T; 25 other;

Query Match 2.5%; Score 37; DB 22; Length 396;
 Best Local Similarity 100.0%; Pred. NO. 7.1e-05;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||
 DB 40 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4

RESULT 14
 AAI88595
 ID AAI88595 standard; cDNA: 411 BP.
 AC
 XX AAI88595;
 XX

DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 8655.
XX
XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
PN
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX MPI: 2001-514838/56.
DR P-PSDB: AAO08664.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1: SEQ ID NO 8655; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 411 BP; 214 A; 24 C; 64 G; 107 T; 2 other;

Query Match 2.5%; Score 37; DB 22; Length 411;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||
DB 156 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 192

RESULT 15
AAI83922
ID AAI83922 standard; cDNA: 412 BP.
XX
XX AAI83922;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 3982.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.

XX
OS Homo sapiens.
XX
XX WO200164835-A2.
PN
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX MPI: 2001-514838/56.
DR P-PSDB: AAO03991.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1: SEQ ID NO 3982; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 412 BP; 161 A; 44 C; 96 G; 104 T; 7 other;

Query Match 2.5%; Score 37; DB 22; Length 412;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||
DB 113 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 149

RESULT 16
AAL19865/c
ID AAL19865 standard; cDNA: 426 BP.
XX
XX AAL19865;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 12322.
DE
XX
XX Human; breast cancer; cell marker; cytoskeletal; ss.
KW
XX
XX Homo sapiens.
OS
XX WO200151628-A2.
PN
PD 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US00798.
PF 14-JAN-2000; 2000US-0176077.
XX 14-MAR-2000; 2000US-0189167.
PR

PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PI Lillie J, Xu Y, Wang Y, Steilmann K.
XX
XX WPI; 2001-451856/48.
DR
XX
XX New peptide useful as a marker for the diagnosis of breast cancer -
PT
PS Claim 1; Page 2177-2178; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AA107544-AA126789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
XX Sequence 426 BP; 125 A; 56 C; 74 G; 171 T; 0 other;
SQ
Query Match 2.5%; Score 37; DB 22; Length 426;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 47 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11
RESULT 17
AAA96571
ID AAA96571 standard; DNA; 524 BP.
XX
AC AAA96571;
XX
XX 08-FEB-2001 (first entry)
DE Noncoding region of core 2 beta-1,6-N-acetylglucosaminyltransferase gene.
XX
XX Core 2 beta-1,6-N-acetylglucosaminyltransferase; core2b GlcNAc-T;
KW cancer; cardiovascular disorder; inflammatory disorder; aschma;
KW rheumatoid arthritis; inflammatory bowel disease; arteriosclerosis;
KW septic shock; adult respiratory distress syndrome; ARDS; cancer;
KW platelet-mediated pathology; atherosclerosis; gastrointestinal disorder;
KW clotting; ascites; cholelithiasis; cirrhosis; Crohn's disease;
KW diverticulitis; ulcerative colitis; ss.
XX
XX Homo sapiens.
OS
XX
XX CA2296936-A1.
PN
XX
XX 03-AUG-2000.
PD
XX
XX 03-FEB-2000; 2000CA-2296936.
PF
XX
XX 03-FEB-1999; 99US-0118674.
PR
XX
XX (GLYC-) GLYCDESIGN INC.
PA
PI Korczak B, Lew A;
XX
XX WPI; 2000-594746/57.
DR
XX
XX New nucleic acid molecules of core 2
PT beta-1,6-N-acetylglucosaminyltransferase useful for providing new

PT compositions for treatment of disorders mediated by the enzyme
PT including cancer, cardiovascular and inflammatory disorders.
XX
XX
PS Disclosure; Page 52; 66pp; English.
XX
XX The present sequence represents a noncoding region from a human core 2
CC beta-1,6-N-acetylglucosaminyltransferase (core2b GlcNAc-T) gene.
CC The polypeptide can be used to treat diseases and disorders, such as
CC cancer, cardiovascular disorders and inflammatory disorders including
CC asthma, rheumatoid arthritis, inflammatory bowel disease,
CC arteriosclerosis, septic shock, adult respiratory distress syndrome
CC (ARDS) and cancer. Various platelet-mediated pathologies such as
CC atherosclerosis and clotting can also be treated. The polypeptides of
CC the invention are predominantly expressed in gastrointestinal tissue
CC (stomach, colon, intestine, testis) and are elevated in cancer.
CC Gastrointestinal disorders that may be prevented or treated include
CC ascites, cholelithiasis, cirrhosis, Crohn's disease, diverticulitis
CC and ulcerative colitis. The antibodies may be used in
CC immuno-histochemical analysis, to detect the novel polypeptide and to
CC localize it to particular cells and tissues and to specific subcellular
CC locations and to quantitate the level of expression.
XX
XX Sequence 524 BP; 171 A; 90 C; 127 G; 136 T; 0 other;
SQ
Query Match 2.5%; Score 37; DB 21; Length 524;
Best Local Similarity 100.0%; Pred. No. 6.7e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 452 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 488
RESULT 18
AAFe4213
ID AAF64213 standard; cDNA; 699 BP.
XX
XX AAF64213;
XX
XX 06-APR-2001 (first entry)
DE Human secreted protein gene 38 SEQ ID NO:48.
XX
XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; vituicide; fungicide;
KW ophthalmological; vulnerrary; autoimmune disease; cardiovascular disorder;
KW hyperproliferative disorders; cerebrovascular disorder; wound healing;
KW nervous system disorder; ocular disorder; skin aging; chemotaxis;
XX
XX Homo sapiens.
OS
XX
XX WO200077026-A1.
PN
XX
XX 21-DEC-2000.
PD
XX
XX 01-JUN-2000; 2000WO-US14973.
PF
XX
XX 11-JUN-1999; 99US-0138630.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX
PI Rosen CA, Ruben SM, Komatsoulis GA;
XX
XX WPI; 2001-071258/08.
DR
XX
XX P-PDB; AAB75543.
DR
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX


```

XX Claim 4: Pages 5-6; 10pp; Japanese.
XX
XX AAQ79736 encodes AAR6604 a flower style-specific S-ribonuclease,
CC isolated from a self-incompatible wild tomato species.
XX
XX Sequence 826 BP; 309 A; 150 C; 131 G; 236 T; 0 other;
SQ
Query Match 2.5%; Score 37; DB 16; Length 826;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 21
AAT64553
ID AAT64553 standard; cDNA to mRNA; 826 BP.
XX
XX AAT64553;
AC
XX
XX 22-MAY-1997 (first entry)
DT
XX
XX Tomato S-ribonuclease cDNA fragment.
DE
XX S-ribonuclease; plant style; promoter; tissue-specific;
KW self-incompatibility; ds.
XX
XX Lycopersicon peruvianum.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..489
FT /tag= a
FT
XX
XX JP09028381-A.
XX
XX 04-FEB-1997.
PD
XX
XX 24-JUL-1995; 95JP-0187557.
PF
XX
XX 24-JUL-1995; 95JP-0187557.
PR
XX
XX (CHEN/) CHENG E. K.
PA (MINA/) MINAMI H.
PA (TAKA/) TAKAGI M.
XX
XX WPI: 1997-159091/15.
DR P-PSDB; AAM14912.
XX
XX Tomato S-ribonuclease gene and promoter sequence - useful for
PT expressing genes in plant style cells, e.g. to confer
PT self-incompatibility on the plant
XX
XX Claim 4; Page 10-11; 15pp; Japanese.
PS
XX The present sequence encodes part of a tomato (Lycopersicon
CC peruvianum) S-ribonuclease. The S-gene can confer self-
CC incompatibility on a plant.
XX
XX Sequence 826 BP; 309 A; 150 C; 131 G; 236 T; 0 other;
SQ
Query Match 2.5%; Score 37; DB 18; Length 826;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 774 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 810

RESULT 22

```

```

AAQ04441
ID AAQ04441 standard; DNA; 957 BP.
XX
XX AAQ04441;
AC
XX 26-SEP-1989 (first entry)
DT
XX
XX Aequorin gene.
DE
XX
XX Photoprotein; aequorin; calcium levels; ss.
KW
XX
XX Aequoria victoria.
OS
XX
XX Key Location/Qualifiers
FH CDS 115..702
FT /tag= a
FT /label=aequorin
FT
XX
XX JP02096597-A.
PN
XX
XX 09-APR-1990.
PD
XX
XX 24-AUG-1984; 84JP-0113000.
PF
XX
XX 01-JAN-1989; 89JP-0176125.
PR
XX
XX (CHRC ) CHITSO CORP.
PA
XX
XX WPI: 1990-151860/20.
DR P-PSDB; AAR04822.
XX
XX Biosynthesis of photoprotein aequorin - comprises culturing modified host
PT conug. specific DNA sequence of photoprotein aequorin
PT
XX
XX Disclosure; ; Japanese.
PS
XX
XX A prokaryotic host is transformed with this sequence in the prodn. of the
CC photoprotein aequorin. The protein is useful for analysing low calcium
CC levels.
CC
XX
XX Sequence 957 BP; 362 A; 156 C; 211 G; 228 T; 0 other;
SQ
Query Match 2.5%; Score 37; DB 11; Length 957;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 890 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 926

RESULT 23
AAC92233
ID AAC92233 standard; cDNA; 958 BP.
XX
XX AAC92233;
AC
XX
XX 21-MAR-2001 (first entry)
DT
XX
XX Apoaequorin-encoding gene nucleotide sequence SEQ ID NO.5.
DE
XX
XX Bioluminescence generating system; fluorescent protein; recreation;
KW entertainment; amusement; toy; paint; slimy play material; textile;
KW bubble making toy; balloon; cosmetic; bath powder; body lotion; gel;
KW powder; cream; toothpaste; dentifrice; soap; body paint; bubble bath;
KW gelatin; icing; frosting; beverage; beer; wine; champagne; soft drink;
KW glowing ice; fountain; liquid firework; ss.
XX
XX Unidentified.
OS
XX
XX US6152358-A.
PN
XX
XX 28-NOV-2000.
PD

```

```

XX 17-AUG-1998; 98US-0135988.
XX
XX 06-FEB-1996; 96US-0597274.
XX
XX 25-NOV-1996; 96US-0757046.
XX
XX (BRYA/) BRYAN B.
XX
XX Bryan B;
XX
XX WPI: 2001-060056/07.
XX
XX P-PSDB: AAB51163.
XX
XX New combinations of articles of manufacture and with bioluminescence
XX generating systems, useful for producing novelty items for
XX entertainment, recreation and amusement, personal items or foods, e.g.
XX toys, cosmetics or gelatins
XX
XX Disclosure: Column 109-112; 83pp; English.
XX
XX
XX The present invention describes a combination comprising an article of
XX manufacture and one or more components of a bioluminescence generating
XX system. The combination is a novelty item and the article of manufacture
XX is a paper product. The combination is useful for producing novelty
XX items, particularly those that glow, or produce or expel a
XX bioluminescent composition. The novelty items are useful for
XX entertainment, recreation and amusement, personal items, or foods.
XX Specifically, these include toys, paints, slimy play materials, textiles
XX (e.g. clothing), bubbles in bubble making toys and other toys that
XX produce bubbles, balloons, cosmetics, bath powders, body lotions, gels,
XX powders and creams, toothpastes and other dentifrices, soaps, body
XX paints, or bubble baths, gelatins, icings and frostings, beverages (beer,
XX wine, champagne, or soft drinks), glowing ice, fountains (e.g. liquid
XX fireworks) and other jets or sprays or aerosols of compositions that are
XX solutions, mixtures, suspensions, powders, pastes, or particles. AAC92229
XX to AAC92241 and AAB51159 to AAB51172 represent sequences which are given
XX in the exemplification of the present invention.
XX
XX Sequence 958 BP; 363 A; 156 C; 210 G; 229 T; 0 other;
XX
XX
XX Query Match 2.5%; Score 37; DB 22; Length 958;
XX Best Local Similarity 100.0%; Pred. No. 6e-05;
XX Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
XX |||||||||||||||||||||||||||||||||||
XX Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927
XX
XX
XX RESULT 24
XX AAD2188
XX ID AAD2188 standard; DNA: 958 BP.
XX
XX AC AAD2188;
XX
XX 12-FEB-2002 (first entry)
XX
XX Aequorea victoria gene.
XX
XX Green fluorescent protein; GFP; bioluminescence generating system; toy;
XX luciferase; finger paint; slimy play material; fishing lure; sparkler;
XX doll; balloon; personal care item; cosmetic; bath powder; body cream;
XX tooth paste; mouth wash; soap; body paint; ornamental transgenic plant;
XX bubble bath; fountain; fairy dust; food; gelatin; icing; frosting; beer;
XX wine; champagne; milk; soft drink; ice cube; ice; foot bag; beverage;
XX dry ice; textile; clothing; paper product; greeting card; wrapping paper;
XX apoequorin; ds.
XX
XX Aequorea victoria.
XX
XX Key Location/Qualifiers
XX CDS 115..705
XX /*tag= a
XX
XX FT

```

```

XX FT
XX /product= "Aequorea victoria apoequorin protein"
XX
XX W0200168824-A2.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-US08277.
XX
XX 15-MAR-2000; 2000US-189691P.
XX
XX (PROL-) PROLUME LTD.
XX (BRYA/) BRYAN B J.
XX
XX Bryan BJ, Szent-Gyorgyi C, Szczepaniak W;
XX
XX WPI: 2002-010561/01.
XX
XX P-PSDB: AAE13370.
XX
XX Nucleic acids encoding Renilla reniformis green fluorescent proteins,
XX useful in diagnostic bioluminescence procedures
XX
XX Disclosure: Page 146-147; 175pp; English.
XX
XX
XX The patent discloses sea pansy (Renilla reniformis) green fluorescent
XX proteins (GFP) and their corresponding polynucleotides. The invention
XX also relates to sequences of the bioluminescence generating system
XX (e.g. luciferase). R. reniformis GFP are used in diagnostic methods
XX and in the production of novelty items such as toys (e.g. squirt gun,
XX pellet gun, toy "Halloween" eggs, toy cigarettes, board/card game
XX toys), finger paints, slimy play material, bubbles in bubble making
XX toys, fishing lures, dolls, sparklers, magic wand toys, balloons,
XX personal care item (e.g. cosmetic, bath powders, body creams, tooth
XX pastes, mouth wash, soaps, body paints, bubble bath), ornamental
XX transgenic plants, fountain, fairy dust, food (gelatins, icings,
XX frosting, beers, wine, champagne, milk, soft drinks, ice cubes, ice,
XX dry ice, beverage), textile (foot bag, clothing) and/or paper product
XX (greeting cards, wrapping paper). The present sequence is a DNA
XX encoding Aequorea victoria apoequorin protein.
XX
XX Sequence 958 BP; 363 A; 156 C; 210 G; 229 T; 0 other;
XX
XX
XX Query Match 2.5%; Score 37; DB 24; Length 958;
XX Best Local Similarity 100.0%; Pred. No. 6e-05;
XX Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
XX |||||||||||||||||||||||||||||||||||
XX Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927
XX
XX
XX RESULT 25
XX AAN81534
XX ID AAN81534 standard; DNA: 959 BP.
XX
XX AC AAN81534;
XX
XX 05-DEC-1990 (first entry)
XX
XX PA0440 aequorin gene variant.
XX
XX PA0440 aequorin; variant; photogenic mechanism; E.coli.
XX
XX Aequorea.
XX
XX Key Location/Qualifiers
XX CDS 220..
XX /*tag= a
XX /*note="G to C"
XX
XX old_sequence 238..
XX /*tag= b
XX /*note="G to A"
XX
XX old_sequence 307..
XX /*tag= c
XX
XX FT

```

```

FT      old_sequence      /note="C to T and also 308 A to T"
FT      499..             /tag= d
FT      /note="G to C"
FT      568..             /tag= e
FT      /note="T to C"
FT      569..             /tag= f
FT      /note="G to C"
FT      590..             /tag= g
FT      /note="G to C"
FT      607..             /tag= h
FT      /note="G to C"
FT      625..             /tag= i
FT      /note="G to A"
FT      616..             /tag= j
FT      /note="G to C and also 625 G to A"
FT      674..             /tag= k
FT      /note="G to C"
FT      205..207          /tag= l
FT      /note="deleted bases GAT"
FT      592..594          /tag= m
FT      /note="deleted bases GAT"
FT      /note="deleted bases GAT"
PN      EP64819-A.
PD      27-APR-1988.
XX
XX      14-OCT-1987;      87EP-0115044.
XX
XX      15-OCT-1986;      86JP-0245108.
XX      23-MAY-1987;      87JP-0126374.
XX
PA      (CHCC ) CHISSO CORPORATION.
XX
PI      Inouye S, Kurose K, Sakaki Y;
XX
DR      WPI; 1988-114006/17.
XX
XX      Variants of aequorin gene - prepd. by site specific mutagenesis process.
XX      Claim 1; page 22; 35pp; English.
XX
CC      The above variant genes are used to prouce variant proteins in E.coli
CC      which enable analysis of the photogenic mechanism of pM0440. Using the
CC      ethanol, an aequorin has been produced which does not require 2-mercapto-
CC      ethanol for its conversion to aequorin.
XX
SQ      Sequence 959 BP; 364 A; 156 C; 210 G; 229 T; 0 other;
XX
Query Match      2.5%; Score 37; DB 9; Length 959;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1427 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
Db      891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

```

```

RESULT 26
AAC77825
ID      AAC77825 standard; cDNA; 1093 BP.
XX
AC      AAC77825;
XX

```

```

DT      08-FEB-2001 (first entry)
XX
DE      Human cancer associated gene sequence SEQ ID NO:219.
XX
KW      Human; cancer associated gene; cancer antigen; detection; cancer;
KW      diagnosis; cytostatic; proliferative; vulneray; immunomodulator;
KW      antidiabetic; antihistaminic; antihemetic; antiarthritic; antiviral;
KW      antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
KW      dermatological; neuroprotective; thrombolytic; coagulant; noctropic;
KW      vasoregic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW      immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW      allergic reaction; graft versus host disease; organ rejection;
KW      haemostatic; thrombolytic; cardiovascular disorder; infection;
KW      neurological disease; drug screening; ss.
XX
OS      Homo sapiens.
XX
PN      WO200055350-A1.
XX
PD      21-SEP-2000.
XX
PF      08-MAR-2000; 2000WO-US05882.
XX
PR      12-MAR-1999; 99US-0124270.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
XX
PI      Rosen CA, Ruben SM;
XX
DR      WPI; 2000-587533/55.
DR      P-PSDB; AAB43616.
XX
PT      Novel isolated nucleic acids comprising sequences encoding peptides
PT      useful for treating or diagnosing e.g. cancer -
XX
PS      Claim 1; Page 789-790; 2352pp; English.
XX
XX
CC      AAC77607 to AAC78448 encode the human cancer associated proteins given
CC      in AAB43398 to AAB44239. The proteins can have activities based on the
CC      tissues and cells the genes are expressed in. Example of activities
CC      include: cytostatic; proliferative; vulnerary; immunomodulator;
CC      antidiabetic; antihistaminic; antihemetic; antiallergic;
CC      antiinflammatory; antihypertoid; antiallergic; antibacterial; antiviral;
CC      dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC      noctropic; vasoregic; antipsoriatic and antiangiogenic. The
CC      polynucleotides and polypeptides can be used for preventing, treating or
CC      ameliorating medical conditions and diagnosing pathological conditions.
CC      Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC      the present invention may be used to treat immune disorders by activating
CC      or inhibiting the proliferation, differentiation or mobilisation of
CC      immune cells, to treat disorders of haematopoietic cells, autoimmune
CC      disorders, allergic reactions, graft versus host disease and organ
CC      rejection, modulate haemostatic or thrombolytic activity, modulate
CC      inflammation, cancers, cardiovascular disorders, neurological disease and
CC      bacterial or viral infections. The peptides, nucleotides, antibodies,
CC      agonists and antagonists may be also be used in drug screens. AAC78449 to
CC      AAC78457 and AAB44240 represent sequences used in the exemplification of
CC      the present invention.
XX
SQ      Sequence 1093 BP; 281 A; 303 C; 330 G; 176 T; 3 other;
XX
Query Match      2.5%; Score 37; DB 21; Length 1093;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1427 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
Db      1014 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1050

```

```

RESULT 27
AAS02547
ID      AAS02547 standard; cDNA; 1159 BP.

```


PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246613.
 PR 08-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI: 2001-465557/50.
 XX
 XX P-PSDB; AAU18211.
 XX
 XX Nucleic acid molecules encoding human secreted chromosomal binding
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
 PT Alzheimer's and Parkinson's diseases and cancers -
 XX
 XX Claim 4; SEQ ID No 68; 561pp; English.
 XX
 XX The present invention relates to the isolation of novel DNA-binding
 CC proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding
 CC for these proteins. DNA-binding proteins such as histones, chromo
 CC (chromatin organisation modifier) domain proteins, and Y-box binding
 CC proteins may contribute to diseases resulting from aberrant DNA
 CC organisation and/or gene transcription. The sequences of the invention
 CC are useful in screening assays to identify antagonists and/or agonists
 CC that may enhance or block activities mediated by DNA-binding proteins.
 CC Blockers of DNA-binding proteins may be useful in treating disorders
 CC such as malignant diseases (e.g. cancer), autoimmune disorders
 CC (e.g. diabetes mellitus), rheumatic diseases (e.g. rheumatoid
 CC arthritis), genetic abnormalities (e.g. cystic fibrosis), infectious
 CC diseases (e.g. HIV) and neurological disorders (e.g. Alzheimer's
 CC disease). The polynucleotide sequences of the invention may also be
 CC used in gene therapy. AAS29030-AAS29157 represent cDNA sequences
 CC encoding for novel DNA-binding proteins.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 1252 BP; 410 A; 221 C; 248 G; 373 T; 0 other;
 SQ
 Query Match 2.5%; Score 37; DB 22; Length 1252;
 Best Local Similarity 100.0%; Pred. No. 5; 8e-05;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 Db 1211 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1247
 RESULT 29
 ID AAS63048 standard; cDNA; 1293 BP.
 AC AAS63048;
 XX
 XX 29-JAN-2002 (first entry)
 DT
 XX
 XX Cell death protective sequence CNI-00727.
 DE
 XX Human, protective sequence; cell death; cerebral oedema; infection;
 KW meningitis; degenerative disease; Alzheimer's disease; heart disease;
 KW motor neuron disease; demyelinating disease; multiple sclerosis; asthma;
 KW nutritional condition; peripheral nervous system disorder; ischemia;
 KW diabetic neuropathy; autoimmune haemolytic anaemia; respiratory system;

KM oral cavity; gastrointestinal tract; liver; cirrhosis; pancreatitis;
 KM polycystic renal disease; urinary tract; genitalia; endometriosis;
 KM breast; chronic mastitis; thyroid gland; Hashimoto's thyroiditis;
 KM adrenal gland; skin; psoriasis; muscular atrophy; bone marrow;
 KM osteoporosis; cancer; autoimmune disease; ss.
 XX
 OS Homo sapiens.
 PN WO200176532-A2.
 PD 18-OCT-2001.
 XX
 XX 09-APR-2001; 2001WO-US11655.
 XX
 XX 11-APR-2000; 2000US-0547596.
 XX
 PA (COGE-) COGENT NEUROSCIENCE INC.
 PI Barney S, Thomas MB, Portbury SD, Putnam K, Katz LC, Lo DC;
 XX
 DR WPI: 2002-017408/02.
 XX
 PT Novel nucleic acids referred as protective sequences and their encoded
 PT products for diagnosing, treating diseases involving cell death,
 PT including neurological disorders e.g. stroke and for identifying
 PT modulators -
 XX
 PS Claim 2; Figure 1E; 256bp; English.
 XX
 CC The invention relates to isolated protective sequence polypeptides (I)
 CC and polynucleotides (II). (II) is useful for transferring a protective
 CC sequence into a cell, which delays and/or prevents the cell from
 CC undergoing cell death. Protective sequences, their products or
 CC antibodies are useful diagnostically, prophylactically, therapeutically
 CC or as targets for treatment and diagnosis of conditions, disorders or
 CC diseases involving cell death. The protective sequences and their
 CC products are useful for preventing or treating disorders of the central
 CC nervous system including neurological and psychiatric conditions,
 CC cerebral oedema, infections such as meningitis, degenerative diseases
 CC such as Alzheimer's and motor neuron disease, demyelinating diseases
 CC as multiple sclerosis, nutritional conditions, disorders of the
 CC peripheral nervous system including diabetic neuropathy, disorders
 CC which cause cell death in organ systems including blood vessels, heart
 CC (ischaemia), blood cells (autoimmune haemolytic anaemia), respiratory
 CC system (asthma), oral cavity; gastrointestinal tract, liver (cirrhosis),
 CC pancreatitis, polycystic renal disease, urinary tract, genitalia
 CC (congenital anomalies), endometriosis, breast (chronic mastitis), thyroid
 CC gland (Hashimoto's thyroiditis), adrenal gland, skin (psoriasis),
 CC musculoskeletal system (muscular atrophy), bone marrow or bone
 CC (osteoporosis). The compositions promote cell death and are useful for
 CC treating and/or ameliorating cancer and autoimmune diseases. The
 CC compounds are further useful for treating physiological impacts on
 CC organs caused by infection which induce cell death. (I) is useful to
 CC raise an immune response, as a reagent in assays designed to
 CC quantitatively determine levels of the protein in biological fluids, as
 CC markers for tissues in which the corresponding protein is expressed and
 CC to isolate receptors or ligands. AAS62967-AAS63142 represent the
 CC protective polynucleotide sequences as described in the invention.
 CC
 XX
 SO Sequence 1293 BP; 484 A; 184 C; 240 G; 385 T; 0 other;

XX
 AC AAV32421;
 XX
 DT 13-OCT-1998 (first entry)
 XX
 DE Homo sapiens clone CC182_1 coding region.
 XX
 KW Secreted protein; clone; CC182_1; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 405..662
 FT CDS /*tag= a
 FT /product= CC182_1 protein
 XX
 PN WO9822501-A2.
 XX
 PD 28-MAR-1998.
 XX
 PF 19-NOV-1997; 97WO-US21123.
 XX
 XX 17-NOV-1997; 97US-0971786.
 PR 20-NOV-1996; 96US-0752912.
 PR 14-FEB-1997; 97US-0800826.
 XX
 BA (GEMV) GENETICS INST INC.
 XX
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racine LA, Spaulding V, Treacy M;
 DR WPI: 1998-312414/27.
 XX
 PS P-PSDB; AAM48807.
 XX
 PT New nucleic acid encoding secreted protein from human cells -
 PT potentially useful, e.g. as immuno-modulators, antitumour agents,
 PT promoters of tissue growth, haemostatic and thrombolytic agents
 XX
 XX Claim 28; Page 70; 93bp; English.
 CC The sequence is that of the coding region of clone CC182_1. It
 CC encodes a secreted protein and may be used to express the protein
 CC recombinantly, as a tissue/molecular weight markers; for chromosome
 CC identification, to identify possible genetic disorders, to isolate
 CC anti-protein or anti-DNA antibodies and in interaction trap
 CC assays to identify sequences that encode interacting proteins. The
 CC protein can be used to screen compounds for biological activity, to
 CC raise antibodies, as tissue markers, for isolation of related receptors
 CC and ligands and as nutritional sources. Such proteins may also
 CC have many biological activities, e.g. cytokine and cell
 CC proliferation/differentiation activity; immunosuppressant or
 CC immunostimulant activity (e.g. for treating immune deficiency, including
 CC infection with human immune deficiency virus, regulation of lymphocyte
 CC growth, treating cancer and many autoimmune diseases, to prevent
 CC transplant rejection or induce tumour immunity), regulation of
 CC haematopoiesis, e.g. treatment of myeloid or lymphoid diseases;
 CC promoting growth of bone, cartilage, tendon, ligament and nerve tissue,
 CC e.g. for healing wounds, treatment of burns, ulcers, periodontal disease
 CC and neurodegeneration, inhibition or activation of follicle-stimulating
 CC hormone (modulation of fertility), chemotactic/chemokine activity
 CC (e.g. for mobilising specific cell types to sites of injury, infection),
 CC haemostatic and thrombolytic activity (e.g. for treating haemophilia or
 CC stroke), as receptors or ligands; anti-inflammatory activity (for treating
 CC septic shock, Crohn's disease etc.), as antimicrobials, modulators of
 CC metabolism and behaviour, as analgesics, enzymes for treating specific
 CC deficiency disorders and in treatment of psoriasis; in human or
 CC veterinary medicine. Neutralising antibodies against the protein can
 CC be used therapeutically, e.g. to detect or prevent metastasis of
 CC cancers expressing the protein. The protein can be expressed in
 CC vivo from DNA, introduced in standard gene therapy vectors.
 XX
 SO Sequence 1362 BP; 409 A; 210 C; 197 G; 544 T; 2 other;

Matches	37	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Qy	1427	AACTAAAAAAAAAAAAAAAAAAAAAAAAA	1463						
Db	1337	AACTAAAAAAAAAAAAAAAAAAAAAAAAA	1373						

RESULT	33
AAAC59580	
ID	AAAC59580 standard; cDNA; 1629 bp.
XX	
XX	
AC	AAAC59580;
XX	
XX	
DT	26-JAN-2001 (first entry)
XX	
XX	
Human secreted protein gene 15 seq	ID NO:25.

KM Human; secreted protein; diagnosis; neuroprotective; cytostatic;
KM cardioactive; immunomodulator; muscular active general; ulnarary;
KM gastrointestinal; nephrotropic; antinefctive; gynaecological; and
KM antibacterial; gene therapy; detection; cancer; chromosome maker;
KM Chromosome identification; neural disorder; immune disorder;
KM muscular disorder; reproductive disorder; gastrointestinal disorder;
KM pulmonary disorder; cardiovascular disorder; renal disorder;
KM proliferative disorder; wound healing; infectious disease; preservative;
KM food additive; ss.

CS	Homo sapiens.
XX	
PN	WO200056883-A1.
XX	
PD	28-SEP-2000.
XX	
PE	16-MAR-2000; 2000MO-US06822.
XX	
PR	23-MAR-1999; 99US-0126054.
PR	10-DEC-1999; 99US-0169916.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsuolis G;
XX
DR WPI; 2000-587666/55.
DR P-PSDB; AAB34313.

PT Human secreted proteins and gene sequences encoding them, useful for
PT detecting, preventing, and treating disorders such as cancer,
PT neurological disorders and immune system disorders -

PS Claim 1; Page 347-348; 429pp; English.

CC The polynucleotide sequences given in AAC59566, AAC59614 encode the
CC human secreted proteins given in AAB34299 to AAB34348 to
CC AAB34437 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular
CC active general; vulnerable; gastrointestinal; nephrotropic;
CC antitumoral; gynaecological; and antibacterial. The polynucleotides
CC can be used for the detection of various disorders such as cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The secreted proteins can be used to
CC treat disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wound healing, and infectious diseases. The proteins can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. AAC59557 to AAC59565 and AAB34298 represent
CC sequences used in the exemplification of the present invention.

Sequence	1629 BP; 477 A; 303 C; 312 G; 537 T; 0 other;
Query Match	2.5%; Score 37; DB 21; Length 1629;

[illegible]

RESULT 34	
AAf63820	
ID	AAf63820 standard; cDNA; 1721 BP.
XX	
AC	AAf63820;
XX	
DT	03-APR-2001 (first entry)
XX	
DE	Human secreted protein gene 32 SEQ ID NO:42.

KM Human, immunosuppressive; antiarthritic; antirheumatic; motropic;
 KM antiproliferative; cytosstatic; cardiant; vasotropic; cerebroprotective;
 KM neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
 KM vulnure; autoimmune disease; hyperproliferative disorder; cancer;
 KM cardiovascular disorder; cerebrovascular disorder; infection;
 KM nervous system disorder; ocular disorder; chemotaxis; food additive;
 KM secreted protein; ss.

OS	Homo sapiens.
XX	
PN	W0200077021-A1.
XX	
PD	21-DEC-2000.
XX	
PF	01-JUN-2000; 2000WO-US15135.
XX	
PR	11-JUN-1999; 99US-0138632.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM, Komatsuolis GA;
XX
DR WP1: 2001-071257/08.
DR P-PSDB: AAB7527L.

PT Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -

PS Claim 1; page 457; 530pp; English.

This invention relates to polynucleotide sequences AAF63789 - AAF63836 which encode human secreted proteins AAB75280 - AAB75287. Included in the invention are protein sequences AAB75288 - AAB75341 which are fragments of the secreted proteins and amino acid sequences with which these fragments share homology. Examples of the activities of the proteins and polynucleotides and the activities of their agonists and antagonists include, immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiatic; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; and vulnery activity. The protein and polynucleotide sequences, their agonists and antagonists may be useful for treating, preventing and diagnosing diseases and disorders such as autoimmune diseases e.g., rheumatoid arthritis, hyperproliferative disorders e.g., neoplasms of the breast or liver, cardiovascular disorders e.g., cardiac arrest, cerebrovascular disorders e.g., cerebral ischaemia, angiodenesis, nervous system disorders e.g., Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. Included in the invention are

CC	oligonucleotides AAF63780 - AAF63788 and peptide AAF5239 which are used
CC	in the identification and characterisation of the DNA and protein
CC	sequences of the invention.
XX	
SQ	Sequence 1721 BP; 558 A; 294 C; 350 G; 519 T; 0 other;
OY	Query Match. 2.5%; Score 37; DB 22; Length 1721;
	Best Local Similarity 100.0%; Pred. No. 5.5e-05;
Db	Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps
OY	1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463 1672 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1708
RESULT 35	
ID	AAD07579 standard; cDNA; 1800 BP.
AC	AAD07579;
XX	
Dt	10-AUG-2001 (first entry)
DE	
XX	Human secreted protein-encoding gene 9 cDNA clone HMWJ52, SEQ ID NO:19
KM	Human; secreted protein; proliferative disorder; cancer; tumour;
KM	fœtal abnormality; developmental abnormality; haematopoietic disorder;
KM	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KM	inflammation; allergy; neurological disorder; Alzheimer's disease;
KM	Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KM	skin disorders; psoriasis; sepsis; diabetes; atherosclerosis;
KM	cardiovascular disorder; angiotensin disorder; kidney disorder;
KM	gastrointestinal disorder; pregnancy-related disorder;
KM	endocrine disorder; infection; wound healing; vulnery;
KM	cell culture; chemotaxis; food additive; gene therapy;
XX	binding partner identification; chromosome 15; ss.
OS	
XX	Homo sapiens.
FH	
FT	Key Location/Qualifiers
FT	CDS 25..180
FT	/tag= a
FT	/product= "Human secreted protein precursor"
FT	sig_peptide 25..96
FT	/tag= b
FT	mat_peptide 97..177
FT	/tag= c
FT	/product= "Mature human secreted protein"
PN	
XX	WO200132676-A1.
XX	
PD	10-MAY-2001.
XX	
PF	25-OCT-2000; 2000WO-US29365.
XX	
PR	29-OCT-1999; 99US-0162237.
XX	21-JUL-2000; 2000US-0219666.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Komatsoulis GA, Shi Y, Olsen HS, Soppet DR:
XX	
DR	WPI: 2001-328773/34.
DR	P-PSTDB; AAEO3060.
XX	
PT	Nucleic acids encoding 25 human secreted polypeptides, useful for
PT	preventing, diagnosing and/or treating e.g. Gaucher's disease,
PT	Alzheimer's disease, Schmitzer syndrome, Creutzfeldt-Jacob disease,
PT	diabetes mellitus and multiple sclerosis -
XX	
PS	Claim 1; Page 415; 546pp; English.
XX	
CC	AAD07571-AAD07645 represent cDNAs corresponding to 25 human secreted

CC	proteinogenes, and AAE03057-AAE03126 represent the proteins they encode.
CC	AAE03127-AAE03150 represent human secreted protein fragments. The genes
CC	and their corresponding secreted proteins are useful for preventing,
CC	treatng or ameliorating medical conditions, e.g., by protein or gene
CC	therapy. Pathological conditions can be diagnosed by determining the
CC	amount of the new protein in a sample or by determining the presence of
CC	mutations in the new genes. Specific uses are described for each of the
CC	25 genes, based on the tissues in which they are most highly expressed,
CC	and include developing products for the diagnosis or treatment of
CC	proliferative disorders, cancer, tumours, foetal and developmental
CC	abnormalities, haematopoietic disorders, diseases of the immune system,
CC	AIDS, autoimmune diseases (e.g., Rheumatoid arthritis), inflammation,
CC	allergies, neurological disorders (e.g., Alzheimer's disease,
CC	Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC	skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC	cardiovascular disorders, angiogenic disorders, kidney disorders,
CC	gastrointestinal disorders, pregnancy-related disorders, endocrine
CC	disorders, and infections. The proteins can also be used to aid wound
CC	healing and epithelial cell proliferation, to prevent skin aging due to
CC	sunburn, to maintain organs before transplantation, for supporting cell
CC	culture of primary tissues, to regenerate tissues, to identify their
CC	cognate ligands or binding partners, and in chemotaxis, and can be used
CC	as a food additive or preservative to modify storage properties.
CC	Antibodies specific for a protein of the invention can be used in
CC	alleviating symptoms associated with the disorders mentioned above, and
CC	in diagnostic immunoassays (e.g., radioimmunoassay or enzyme linked
CC	immunosorbent assay (ELISA). The present sequence represents a human
CC	secreted protein-encoding cDNA of the invention.
XX	
SO	Sequence 1800 BP; 483 A; 403 C; 440 G; 474 T; 0 other:
	Query Match 2.5%; Score 37; DB 22; Length 1800;
	Best Local Similarity 100.0%; Pred. No. 5.4e-05;
	Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1427 AACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Dd	1698 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1734
RESULT 36	
AAA96570	
ID	AAA96570 standard; DNA; 2108 BP.
XX	
AC	AAA96570;
XX	
DT	
XX	
DE	08-FEB-2001 (first entry)
XX	
Core 2 beta-1,6-N-acetylglucosaminyltransferase DNA fragment.	
KW	Core 2 beta-1,6-N-acetylglucosaminyltransferase; core2b glcNAc-T;
KW	cancer; cardiovascular disorder; inflammatory disorder; asthma;
KW	Rheumatoid arthritis; inflammatory bowel disease; arteriosclerosis;
KW	septic shock; adult respiratory distress syndrome; ARDS; cancer;
KW	platelet-mediated pathology; atherosclerosis; gastrointestinal disorder;
KW	clotting; ascites; cholelithiasis; cirrhosis; Crohn's disease;
KW	diverticulitis; ulcerative colitis; ss.
XX	
OS	Homo sapiens.
XX	
PN	CA2296936-AI.
XX	
PD	03-AUG-2000.
XX	
PF	03-FEB-2000; 2000CA-2296936.
XX	
PR	03-FEB-1999; 99US-0118674.
XX	
PA	(GLYC-) GLYCDESIGN INC.
XX	
PI	Korczak B, Lew A;
XX	
DR	WPL; 2000-594746/57.

```
XX PH New nucleic acid molecules of core 2
PT beta-1,6-N-acetylglucosaminyltransferase useful for providing new
PT compositions for treatment of disorders mediated by the enzyme
PT including cancer, cardiovascular and inflammatory disorders.
XX
PS Claim 4: Page 51-52; 66pp; English.
XX
CC The present sequence encodes a partial human core 2
CC beta-1,6-N-acetylglucosaminyltransferase (core2b G1CNC-r) polypeptide.
CC The polypeptide can be used to treat diseases and disorders, such as
CC cancer, cardiovascular disorders and inflammatory disorders including
CC asthma, rheumatoid arthritis, inflammatory bowel disease,
CC arteriosclerosis, septic shock, adult respiratory distress syndrome
CC (ARDS) and cancer. Various platelet-mediated pathologies such as
CC atherosclerosis and clotting can also be treated. The polypeptides of
CC the invention are predominantly expressed in gastrointestinal tissue
CC (stomach, colon, intestine, testis) and are elevated in cancer.
CC Gastrointestinal disorders that may be prevented or treated include
CC ascites, cholelithiasis, cirrhosis, Crohn's disease, diverticulitis
CC and ulcerative colitis. The antibodies may be used in
CC immuno-histochemical analysis, to detect the novel polypeptide and to
CC localize it to particular cells and tissues and to specific subcellular
CC locations and to quantitate the level of expression.
SQ Sequence 2108 BP; 596 A; 464 C; 509 G; 539 T; 0 other;
Query Match
Best Local Similarity 2.5%; Score 37; DB 21; Length 2108;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2036 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2072
RESULT 37
AAA23424
ID AAA23424 standard; cDNA; 2198 BP.
XX
AC AAA23424;
XX
DT 19-JUN-2000 (first entry)
XX
DE cDNA encoding human secreted protein vb12.1, SEQ ID NO:3.
XX
KW Human; secreted protein; cancer; tumour; cardiovascular disorder;
KW blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
KW infection; fungal; bacterial; viral; HIV; allergy; arthritis;
KW neurodegenerative disease; asthma; contraceptive; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 63..485
FH FT /*tag= a
FT /product= "Human secreted protein vb12.1"
XX
PD WO200011015-A1.
XX
PD 02-MAR-2000.
XX
PD 24-AUG-1999; 99WO-US19351.
XX
PF 24-AUG-1999;
XX
PR 24-AUG-1998; 98US-0097638.
XX
PR 24-AUG-1998; 98US-0097659.
XX
PR 09-SEP-1998; 98US-0099618.
XX
PR 28-SEP-1998; 98US-0102092.
XX
PR 25-NOV-1998; 98US-0109978.
XX
PR 23-DEC-1998; 98US-0113645.
XX
PR 23-DEC-1998; 98US-0113646.
XX
PR 23-AUG-1999; 99US-0379246.
XX
```

```
PA (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P.
XX WPI: 2000-224657/19.
DR P-PSDB; AAY94982.
XX
PT New secreted or transmembrane proteins and polynucleotides encoding
PT them, useful for treating neurodegenerative disorders, autoimmune
PT diseases and cancer.
XX
PS Claim 12: Page 264-265; 357pp; English.
XX
CC The invention relates to 40 human secreted proteins (AAY94981-Y95020),
CC and cDNA sequences encoding them (AAA23423-A23462). The secreted
CC proteins of the invention include those that are thought to be only
CC partially secreted, i.e., transmembrane proteins. The proteins of the
CC invention may exhibit one or more activities selected from the following:
CC cytokine activity; cell proliferation; differentiation; immune
CC modulation; haematopoiesis regulation; tissue growth activity;
CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
CC and thrombolytic activity; anti-inflammatory activity; and tumour
CC inhibition activity. The proteins may be administered to patients as
CC vaccines, and the nucleotides may be used as part of a gene therapy
CC regime. Diseases or conditions that may be treated using the proteins or
CC nucleotides of the invention include autoimmune diseases; genetic
CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
CC fungal and viral infections, especially HIV; multiple sclerosis;
CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
CC insulin dependent diabetes mellitus; and allergic reactions such as
CC asthma and anaemia. They may also be used for treating wounds, burns,
CC disease, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
CC additionally be useful as contraceptives. Nucleic acid sequences of the
CC invention may be used in chromosome mapping, and as a source of
CC diagnostic primers and probes. The present sequence represents cDNA
CC encoding one of the 40 proteins of the invention.
XX
SQ Sequence 2198 BP; 618 A; 460 C; 531 G; 589 T; 0 other;
Query Match
Best Local Similarity 2.5%; Score 37; DB 21; Length 2198;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2087 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2123
RESULT 38
AAI97914
ID AAI97914 standard; cDNA; 2260 BP.
XX
AC AAI97914;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 3899.
XX
KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS Homo sapiens.
XX
PD WO200166719-A1.
XX
PD 13-SEP-2001.
XX
PD 02-MAR-2001; 2001WO-JP01629.
XX
PF 07-MAR-2000; 2000JP-0159195.
XX
PR (CHIB-) CHIBA PREFECTURE.
XX
```

PA (HISM) HISAMITSU PHARM CO LTD.
XX
PI Nakagawara A;
XX
DR WPI: 2001-565584/63.
XX
PT Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT for anti-cancer agents -
XX
PS Claim 1; Page 2935-2936; 2979pp; Japanese.
XX
CC The invention relates to novel genes (AA193926-AA197963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes.
XX
SQ Sequence 2260 BP; 677 A; 503 C; 509 G; 571 T; 0 other;

Query Match 2.5%; Score 37; DB 22; Length 2260;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2152 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2188

RESULT 39
AA198068
ID AA198068 standard; cDNA; 2260 BP.
XX
AC AA198068;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 11.
XX
KW Human; neuroblastoma; ss.
XX
OS Homo sapiens.
XX
PN WO200166733-A1.
XX
PD 13-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-JP01631.
XX
PR 07-MAR-2000; 2000JP-0159195.
PR 12-MAY-2000; 2000JP-0140387.
XX
PA (CHIB-) CHIBA PREFECTURE.
PA (HISM) HISAMITSU PHARM CO LTD.
PI Nakagawara A;
XX
DR WPI: 2001-602630/68.
XX
PT Nucleic acids for prognosis of human neuroblastoma comprise nucleic
PT acids expressed by human neuroblastomas -
XX
PS Claim 1; Page 51-53; 159pp; Japanese.
XX
CC The invention relates to nucleic acids (AA198058-AA199161) or their
CC homologues expressed by human neuroblastomas useful for detecting genes
CC expressed by neuroblastoma and for analysing their structure and
CC function. The nucleic acids are useful for the diagnosis and prognosis of
CC neuroblastoma.
XX
SQ Sequence 2260 BP; 677 A; 503 C; 509 G; 571 T; 0 other;

Query Match 2.5%; Score 37; DB 22; Length 2260;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2152 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2188

RESULT 40
AA293309
ID AA293309 standard; DNA; 2297 BP.
XX
AC AA293309;
XX
DT 04-JUL-2000 (first entry)
XX
DE Sequence encoding mouse Homer-3.
XX
DE Homer; calcium; receptor; immediate early gene; IEG;
KW identification; treatment; glutamate receptor;
KW inositol triphosphate; epilepsy; glutamate toxicity;
KW memory disorder; learning disorder; stroke; schizophrenia;
KW Alzheimer's disease; tissue degeneration; brain development;
KW cardiac disorder; muscular disorder; vascular disorder;
KW neurological disorder; psychiatric disorder; renal disorder;
KW uterine disorder; bronchial disorder; ageing; mouse; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FH 1..1177
FT CDS /*tag= a
FT /product= Homer-3
XX
FN WO200011204-A2.
XX
PD 02-MAR-2000.
XX
PF 18-AUG-1999; 99WO-US18973.
XX
PR 18-AUG-1998; 98US-0097334.
PR 09-JUN-1999; 99US-0138426.
PR 09-JUN-1999; 99US-0138493.
PR 09-JUN-1999; 99US-0138494.
XX
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
PI Worley PF, Tu JC, Xiao B, Leahy D, Beneken J, Lanhahan AA;
XX
DR WPI: 2000-246571/21.
DR P-PSDB; AAY83013.
XX
PT Identifying compounds capable of modulating cellular response useful
PT for treating Alzheimer's disease and cardiac disorders, involves
PT incubating compound with cell expressing Homer protein and cell surface
PT receptor -
XX
PS Example 1; Page 151; 171pp; English.
XX
CC Homer proteins are the products of neuronal immediate early genes
CC (IEG's). They selectively bind the carboxy terminl of certain
CC cell-surface receptors, certain intracellular receptors and binding
CC proteins. Many forms of Homer proteins contain a "coiled-coil"
CC structure in the carboxy terminal domain which mediated homo- and
CC heteromultimerisation between Homer proteins. Homer plays a
CC significant role in mediating receptor-activated calcium mobilisation
CC from intracellular stores. Thus, cells expressing a Homer protein
CC can be used to identify a compound capable of modulating a cellular
CC response mediated by cell surface receptor or intracellular receptor.
CC Compounds identified in this manner which modulate Homer protein
CC activity are useful for treating disorders associated with glutamate

PS Claim 9; Page 124-125; 131pp; English.

CC fragment is useful for promoting bone growth, treating bone fractures,

CC prophylactically increasing or maintaining bone density in a subject
CC having a substantially normal bone density, so as to stimulate
CC osteoblast proliferation, treating osteoporosis, muscle or bone loss
CC due to malignancy, endocrine disorder, arthritis, sarcopena and
CC periodontal disease, preventing cartilage differentiation, and
CC wound healing.

CC Sequence 2827 BP; 864 A; 522 C; 589 G; 852 T; 0 other;

Query Match 2.5%; Score 37; DB 22; Length 2827;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2788 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2824

RESULT 43

AAV40744 ID AAV40744 standard; cDNA; 2836 BP.

AAV40744;

23-SEP-1998 (first entry)

C. felis esterase, nFE72836, coding sequence.

Esterase; flea; protective immune response; carboxylesterase; arthropod;
haematophagous ectoparasite infestation; nFE72836; ds.

Ctenocephalides felis.

Key Location/Qualifiers

CDS 99..1889

WO9821324-A1.

22-MAY-1998.

10-NOV-1997; 97WO-US20598.

12-NOV-1996; 96US-0747221.

(HESK-) HESKA CORP.

Brandt KS, Silver GM, Wisniewski N;

WPI: 1998-297929/26.

P-PSDB; AAM57856.

New nucleic acid encoding carboxylesterase(s) from fleas - useful,
e.g. in vaccines, for preventing infestation by haematophagous
ectoparasites, particularly on cats and dogs

Claim 1; Page 129-133; 230pp; English.

XX This sequence encodes the flea esterase protein, nFE72836 (the
CC complementary strand is shown in AAV40745), of the invention. When
CC administered to animals, the protein induces a protective immune
CC (antibody) response against carboxylesterase (CE), so they, or
CC compositions containing CE proteins, are used, therapeutically or as
CC vaccines, to protect particularly mammals and birds, specifically cats
CC and dogs, against haematophagous ectoparasite infestation (HEP).
CC Specifically fleas. More generally the compositions can be used to treat
CC arthropods generally, including pests of agricultural crops, trees,
CC stored goods etc., also those that are vectors of disease. Fragments of
CC the DNA can be used as probes and primers for identification or
CC production of nucleic acid. Antibodies against the protein can be used
CC for passive immunisation, to screen expression libraries: to isolate the
CC protein and to target cytotoxic compounds to HEP. The compounds
CC containing CE are effective against both adult and larval stages: they

CC target CEs, including juvenile hormone, that are involved in
CC development, metamorphosis, feeding, digestion and reproduction.

XX Sequence 2836 BP; 1064 A; 421 C; 473 G; 876 T; 2 other;

Query Match 2.5%; Score 37; DB 19; Length 2836;

Best Local Similarity 100.0%; Pred. No. 5e-05;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463

Db 2780 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2816

RESULT 44

AAV40745/C ID AAV40745 standard; cDNA; 2836 BP.

AAV40745;

23-SEP-1998 (first entry)

C. felis esterase, nFE72836, coding sequence complementary strand.

Esterase; flea; protective immune response; carboxylesterase; arthropod;
haematophagous ectoparasite infestation; nFE72836; ds.

Ctenocephalides felis.

WO9821324-A1.

22-MAY-1998.

10-NOV-1997; 97WO-US20598.

12-NOV-1996; 96US-0747221.

(HESK-) HESKA CORP.

Brandt KS, Silver GM, Wisniewski N;

WPI: 1998-297929/26.

P-PSDB; AAM57856.

New nucleic acid encoding carboxylesterase(s) from fleas - useful,
e.g. in vaccines, for preventing infestation by haematophagous
ectoparasites, particularly on cats and dogs

Claim 1; Page 136-137; 230pp; English.

XX This sequence is the complementary strand of the DNA encoding the flea
CC esterase protein, nFE72836 (see AAV40744 for coding strand), of the
CC invention. When administered to animals, the protein induces a
CC protective immune (antibody) response against carboxylesterase (CE), so
CC they, or compositions containing CE proteins, are used, therapeutically
CC or as vaccines, to protect particularly mammals and birds, specifically
CC cats and dogs, against haematophagous ectoparasite infestation (HEP).
CC Specifically fleas. More generally the compositions can be used to treat
CC arthropods generally, including pests of agricultural crops, trees,
CC stored goods etc., also those that are vectors of disease. Fragments of
CC the DNA can be used as probes and primers for identification or
CC production of nucleic acid. Antibodies against the protein can be used
CC for passive immunisation, to screen expression libraries: to isolate the
CC protein and to target cytotoxic compounds to HEP. The compounds
CC containing CE are effective against both adult and larval stages: they
CC target CEs, including juvenile hormone, that are involved in
CC development, metamorphosis, feeding, digestion and reproduction.

XX Sequence 2836 BP; 877 A; 473 C; 421 G; 1064 T; 1 other;

Query Match 2.5%; Score 37; DB 19; Length 2836;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Job time : 261.666 secs

OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 |||||||||||||||||||||||||||||||||||
 DB 56 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20

RESULT 45

AAD21170
 ID AAD21170 standard; CDNA; 2836 BP.

XX
 AC AAD21170;

XX
 DT 15-JAN-2002 (first entry)

XX
 DE Ctenocephalides felis carboxylesterase full length cDNA, nfe72836.

XX
 KM Carboxylesterase; haematophagous arthropod infestation; protozoa;

XX
 KW flea allergic dermatitis; nematode; cestode; trematode; vaccine;

XX
 OS Ctenocephalides felis.

XX
 FH Key Location/Qualifiers

FT CDS 99..1889

FT /*tag= a

FT /product= "Carboxylesterase full length protein, PFE7596"

FT /note= "This region is specifically claimed in claim 1

FT as SEQ ID NO:28"

XX
 PN US6291222-B1.

XX
 PD 18-SEP-2001.

XX
 PF 09-JAN-1998; 98US-0005051.

XX
 PR 09-JAN-1998; 98US-0005051.

XX
 PA (HESK-) HESKA CORP.

XX
 PI Silver GM, Wisniewski N;

XX
 DR WPI: 2001-647226/74.

XX
 DR P-PSDB; AAE12912.

XX
 PT New carboxylesterase nucleic acids and proteins useful for protecting

XX
 PT animals from hemaphysal arthropod infestation, which cause diseases

XX
 PT (e.g. flea allergic dermatitis) and carry infectious agents (e.g.

XX
 PT nematodes)

XX
 PS Claim 1: Column 103-110; 109pp; English.

XX
 CC The invention relates to arthropod such as Ctenocephalides felis

XX
 CC (flea) carboxylesterase (CE) proteins and nucleic acid molecules

XX
 CC encoding such proteins. Sequences of the invention are particularly

XX
 CC useful for protecting animals (e.g. humans, dogs, cattle or zoo animals)

XX
 CC from haematophagous arthropod infestation which cause diseases such

XX
 CC as flea allergic dermatitis (FAD) and carry infectious agents (e.g.

XX
 CC nematodes, cestodes, trematodes or protozoa). They are also useful

XX
 CC in gene therapy and as vaccines. The present sequence is C. felis

XX
 CC carboxylesterase full length cDNA referred to as nfe72836.

XX
 SQ Sequence 2836 BP: 1064 A; 421 C; 473 G; 877 T; 1 other;

XX
 Query Match 2.5%; Score 37; DB 22; Length 2836;

XX
 Best Local Similarity 100.0%; Pred. No. 5e-05;

XX
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
 OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463

XX
 DB 2781 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2817

Search completed: November 5, 2002, 13:46:49

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 11:01:37 ; Search time 35.881 Seconds
(without alignments)
10015.359 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463
Sequence: 1 caccgagaatagctcgcgcc.....aaaaaaaaaaaaaaaa 1463

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 38353 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PC/US.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1463	100.0	1463	US-09-426-557-1	Sequence 1, Appli
2	1360	93.0	1541	US-09-426-557-3	Sequence 3, Appli
3	1267	86.6	1381	US-09-426-557-5	Sequence 3, Appli
4	894	61.1	1478	US-09-426-557-7	Sequence 7, Appli
5	37	2.5	958	US-08-757-046A-5	Sequence 5, Appli
6	37	2.5	958	US-09-447-208-5	Sequence 5, Appli
7	37	2.5	958	US-09-135-988-5	Sequence 5, Appli
8	37	2.5	958	US-09-277-716-5	Sequence 5, Appli
9	37	2.5	958	US-08-597-274A-5	Sequence 5, Appli
10	37	2.5	2836	US-08-747-221B-24	Sequence 24, Appli
11	37	2.5	2836	US-08-747-221B-26	Sequence 24, Appli
12	37	2.5	2836	US-09-005-051-24	Sequence 26, Appli
13	37	2.5	2836	US-09-005-051-26	Sequence 26, Appli
14	36	2.5	2237	US-08-487-135B-1	Sequence 26, Appli
15	36	2.5	2237	US-08-915-672A-1	Sequence 1, Appli
16	35	2.5	2237	US-09-177-909-1	Sequence 1, Appli
17	35	2.4	340	US-08-171-385-27	Sequence 27, Appli
18	35	2.4	340	US-08-361-441B-27	Sequence 27, Appli
19	35	2.4	347	US-08-104-072B-2	Sequence 2, Appli
20	35	2.4	350	US-08-171-385-14	Sequence 14, Appli
21	35	2.4	350	US-08-361-441B-14	Sequence 14, Appli
22	35	2.4	593	US-09-385-982-262	Sequence 16, Appli
23	35	2.4	740	US-08-713-000-8	Sequence 8, Appli
24	35	2.4	740	US-08-975-316-8	Sequence 8, Appli
25	35	2.4	740	US-09-211-710-8	Sequence 8, Appli
26	35	2.4	741	US-08-975-316-8	Sequence 58, Appli
27	35	2.4	1075	US-08-400-006B-6	Sequence 6, Appli

28	35	2.4	1265	US-08-991-789A-169	Sequence 169, App
29	35	2.4	1265	US-09-062-451-169	Sequence 169, App
30	35	2.4	1315	US-09-721-822A-10	Sequence 10, Appli
31	35	2.4	1483	US-09-262-749-1	Sequence 1, Appli
32	35	2.4	1522	US-09-413-574-1	Sequence 1, Appli
33	35	2.4	1551	US-09-461-474-7	Sequence 7, Appli
34	35	2.4	1553	US-09-022-669-1	Sequence 1, Appli
35	35	2.4	1609	US-08-976-808C-2	Sequence 2, Appli
36	35	2.4	1637	US-08-852-824-3	Sequence 3, Appli
37	35	2.4	1639	US-08-333-358-7	Sequence 7, Appli
38	35	2.4	1659	US-08-463-694-7	Sequence 7, Appli
39	35	2.4	1659	US-08-694-501-7	Sequence 7, Appli
40	35	2.4	1728	US-08-985-950-7	Sequence 7, Appli
41	35	2.4	1776	US-08-655-352-10	Sequence 10, Appli
42	35	2.4	1824	US-08-606-505B-1	Sequence 1, Appli
43	35	2.4	1824	US-09-616-990-1	Sequence 1, Appli
44	35	2.4	2399	US-09-385-801-1	Sequence 1, Appli
45	35	2.4	2502	US-09-234-332-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-426-557-1
Sequence 1, Application US/09426557
Patent No. 6232527
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
FILE REFERENCE: 0961
CURRENT APPLICATION NUMBER: US/09/426,557
CURRENT FILING DATE: 1999-10-22
EARLIER APPLICATION NUMBER: 60/112,332
EARLIER FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1463
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (85)...(1221)
US-09-426-557-1

Query Match 100.0%; Score 1463; DB 4; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CACGAGATAGCTCGCGCGCGGTTCTTCTGCGCATCCGCGTCACGCCCGCCGCCA	60
DB	1	CACGAGATAGCTCGCGCGCGGTTCTTCTGCGCATCCGCGTCACGCCCGCCGCCA	60
QY	61	CCCCCAGCGCGCGCGAGAGATGGCATCAAGGTTTGAAGAACTGTCGGCGAC	120
DB	61	CCCCCAGCGCGCGCGAGAGATGGCATCAAGGTTTGAAGAACTGTCGGCGAC	120
QY	121	AATCGCCCAAGCGCGATGAAGAGACAGAGTTGAGAGCTACTTGGCCGCAAAATCGCC	180
DB	121	AATCGCCCAAGCGCGATGAAGAGACAGAGTTGAGAGCTACTTGGCCGCAAAATCGCC	180
QY	181	GTGCGCGCGAGATGACATATACCACTTCTGATATGTTGAAGAGACGATGGAA	240
DB	181	GTGCGCGCGAGATGACATATACCACTTCTGATATGTTGAAGAGACGATGGAA	240
QY	241	ACTCTCAAAATGAAGCTGTGAAGTCACTAGTATTTGCAAGAAATGTTCAACCGGACA	300
DB	241	ACTCTCAAAATGAAGCTGTGAAGTCACTAGTATTTGCAAGAAATGTTCAACCGGACA	300
QY	301	ATAAGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTGTATGGCAAGCCTCTGTAT	360
DB	301	ATAAGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTGTATGGCAAGCCTCTGTAT	360

Db 301 ATAGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTTGATGGCAAGCCTCTGAT 360
QY 361 ATGAACAACAAGAGCTTCTAAAGATCTCAAAAAGATGATGCAAC*AAAGATCG 420
Db 361 ATGAAGAACAGAGCTTCTTAAGATCTCAAAAAGATGATGCAACCAAAATCTG 420
QY 421 ACTGAGCAGTAGAGGTAGAGATTAAGATGCGATTGAAAAATTGACAAGAGACTGTA 480
Db 421 ACTGAGCAGTAGAGGTAGAGATTAAGATGCGATTGAAAAATTGACAAGAGACTGTA 480
QY 481 AAGGTACACAAGGACACACAAGATTTGTAACGCTATTAAAGCTTATGGGGTCTCT 540
Db 481 AAGGTACACAAGGACACACAAGATTTGTAACGCTATTAAAGCTTATGGGGTCTCT 540
QY 541 GTTGAGAGGACCTCTCTAAGCAGAGAGATGCGACCTTGTGCTAAACATAG 600
Db 541 GTTGAGAGGACCTCTCTAAGCAGAGAGATGCGACCTTGTGCTAAACATAG 600
QY 601 GTGTGCGTGTGCTTCAAGATATGACTCCCTTACTTTTGGGGCTCCACGTTCTT 660
Db 601 GTGTGCGTGTGCTTCAAGATATGACTCCCTTACTTTTGGGGCTCCACGTTCTT 660
QY 661 CGTCATTTAATGATGCCAAGTCCAGAAATATACCTGTGATGGAATTTGATGCCAAG 720
Db 661 CGTCATTTAATGATGCCAAGTCCAGAAATATACCTGTGATGGAATTTGATGCCAAG 720
QY 721 GTTTGAGAGAGCTTGAACCTACCATGAGCACAGTTCATTGATTGTCATCTGTGGA 780
Db 721 GTTTGAGAGAGCTTGAACCTACCATGAGCACAGTTCATTGATTGTCATCTGTGGA 780
QY 781 TGTGACTATTTGATAGCATCAAAAGTATCGGGGGCAACAGCTTGAACCTTATTCGT 840
Db 781 TGTGACTATTTGATAGCATCAAAAGTATCGGGGGCAACAGCTTGAACCTTATTCGT 840
QY 841 CAACATGGGCTCATAGAAAGCATCTTGGAATCTTAATAAAGACAGATCAAAATTCCT 900
Db 841 CAACATGGGCTCATAGAAAGCATCTTGGAATCTTAATAAAGACAGATCAAAATTCCT 900
QY 901 GAGAGACTGGCCTTACCAAGAGCTCGACGCTTGTTCAAGAGAGCTTAATGTCAATTGAT 960
Db 901 GAGAGACTGGCCTTACCAAGAGCTCGACGCTTGTTCAAGAGAGCTTAATGTCAATTGAT 960
QY 961 ATTCTGAGCTAAATAGAGTGCACCTGATGAGAGGGTCTCAATAGTTTCTGTGTA 1020
Db 961 ATTCTGAGCTAAATAGAGTGCACCTGATGAGAGGGTCTCAATAGTTTCTGTGTA 1020
QY 1021 GATAATGCTTCAACGAGATCGGTGACAAAGCCATAGAGAGATCAAAATCTGCCAAG 1080
Db 1021 GATAATGCTTCAACGAGATCGGTGACAAAGCCATAGAGAGATCAAAATCTGCCAAG 1080
QY 1081 AATAAATCGTGCAGAGAGACTCGAGTCTTTTCAAGCCAACTGCCACACATCAGCA 1140
Db 1081 AATAAATCGTGCAGAGAGACTCGAGTCTTTTCAAGCCAACTGCCACACATCAGCA 1140
QY 1141 CCGCTAAAGGGAAGGAGACTTGGAATAAACAAGCAAGGACCTCGCAACAGAAACA 1200
Db 1141 CCGCTAAAGGGAAGGAGACTTGGAATAAACAAGCAAGGACCTCGCAACAGAAACA 1200
QY 1201 AAGGCTGTGGAAGAAGAAATATCTTGGATGCTTGAATACATAGACTAGCAAG 1260
Db 1201 AAGGCTGTGGAAGAAGAAATATCTTGGATGCTTGAATACATAGACTAGCAAG 1260
QY 1261 CAGCGTGGCGTATCCTTGCCTTATTTAACTCCCTGTTTAACTCAGAGCTTT 1320
Db 1261 CAGCGTGGCGTATCCTTGCCTTATTTAACTCCCTGTTTAACTCAGAGCTTT 1320
QY 1321 GGTAAAGTGTGCTCAAGTTCGAGTGGGTAGTGTGTTGAAGACATTTGGTG 1380
Db 1321 GGTAAAGTGTGCTCAAGTTCGAGTGGGTAGTGTGTTGAAGACATTTGGTG 1380
QY 1381 TACCAAGTAAACAACTTATCGCTGTTTTTACTTCTGTGCTTTGAAGTAAAAA 1440
Db 1381 TACCAAGTAAACAACTTATCGCTGTTTTTACTTCTGTGCTTTGAAGTAAAAA 1440

QY 1441 AAAAAAAAAAAAAAAAAAAAA 1463
Db 1441 AAAAAAAAAAAAAAAAAAAAA 1463

RESULT 2
US-09-426-557-3
; Sequence 3, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)...(1215)
US-09-426-557-3

Query Match 93.08; Score 1360; DB 4; Length 1541;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 GCGGTTTCTGCGCACATCCGGCTCAGCCGCCGCCCAACCCGCCACAGCCCGCAGA 80
Db 15 GCGGTTTCTGCGCACATCCGGCTCAGCCGCCGCCCAACCCGCCACAGCCCGCAGA 74
QY 81 CGAGATGGGCATCAAGGTTTGACGAAATCTGTCGGGAGACATGCGCCCAAGCGATGA 140
Db 75 CGAGATGGGCATCAAGGTTTGACGAAATCTGTCGGGAGACATGCGCCCAAGCGATGA 134
QY 141 GGAGAGAGATTCGAGAGCTACTCGGCCGCAAAATGCGGTGAGCGCAGCATGAGCAT 200
Db 135 GGAGAGAGATTCGAGAGCTACTCGGCCGCAAAATGCGGTGAGCGCAGCATGAGCAT 194
QY 201 ATACAGTTCCTGATTTGATTTGGAAGACAGCATGGAATCTCTCAAAATGAAGCTGG 260
Db 195 ATACAGTTCCTGATTTGATTTGGAAGACAGCATGGAATCTCTCAAAATGAAGCTGG 254
QY 261 TGAAGTCACTAGTCATTTGCAAGGATGTTCAACCGGACATTAAGATTACTGGAAGCGG 320
Db 255 TGAAGTCACTAGTCATTTGCAAGGATGTTCAACCGGACATTAAGATTACTGGAAGCGG 314
QY 321 AATCAGCCAGTTATGTTTGTGATGGCAAGCTCTGATTTGGAAGAACAAGAGCTTGC 380
Db 315 AATCAGCCAGTTATGTTTGTGATGGCAAGCTCTGATTTGGAAGAACAAGAGCTTGC 374
QY 381 TAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTGACTGAGGACAGTAGAGTAGG 440
Db 375 TAAAGATACCTCAAAAGAGATGATGCAACCAAGATCTGACTGAGGACAGTAGAGTAGG 434
QY 441 AGATTAAGATGCGATTTGAAAAATTGAGCAGAGAGACGTTAAAGTTCACAAAGGACACAA 500
Db 435 AGATTAAGATGCGATTTGAAAAATTGAGCAGAGAGACGTTAAAGTTCACAAAGGACACAA 494
QY 501 CGAAGATTGTAAGCGCATTAAGACTATGAGGAGGTTCTGTTTGAAGGACACTTCTGA 560
Db 495 CGAAGATTGTAAGCGCATTAAGACTATGAGGAGGTTCTGTTTGAAGGACACTTCTGA 554
QY 561 AGCAGAGCAGATGTCAGGCCCTTTGCAATAAGCATAAAGTGTTCCTGTTGCTTCA 620
Db 555 AGCAGAGCAGATGTCAGGCCCTTTGCAATAAGCATAAAGTGTTCCTGTTGCTTCA 614

```
QY 621 AGATATGACCTCCCTTACTTTTGGGGCTCCACGGTTCCTGTCATTTAATGATCCAG 680
    |||
Db 615 AGATATGACCTCCCTTACTTTTGGGGCTCCACGGTTCCTGTCATTTAATGATCCAG 674
QY 681 TTCAGAAAATACCTGTGATGGAATTTGATGTGGCAAGTTTGGAGAGCTTGAAT 740
    |||
Db 675 TTCAGAAAATACCTGTGATGGAATTTGATGTGGCAAGTTTGGAGAGCTTGAAT 734
QY 741 CACCATGAGACCTTCATGATTTGTGCAATCCGTGTGATGTGACTATTTGTATGACAT 800
    |||
Db 735 CACCATGAGACCTTCATGATTTGTGCAATCCGTGTGATGTGACTATTTGTATGACAT 794
QY 801 CAAAGGTATCGGGGGGCAAAACAGCTCTGAAACTTATTCGCAACATGGGTCATAGAAAG 860
    |||
Db 795 CAAAGGTATCGGGGGGCAAAACAGCTCTGAAACTTATTCGCAACATGGGTCATAGAAAG 854
QY 861 CATCTTGGAGAACTCTTAATTAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGA 930
    |||
Db 855 CATCTTGGAGAACTCTTAATTAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGA 914
QY 921 ACCTGACGCTTGTTCAGAGAGCTTAATGTCACATGATGATATTCGAGGTAAATAGAGAC 980
    |||
Db 915 ACCTGACGCTTGTTCAGAGAGCTTAATGTCACATGATGATATTCGAGGTAAATAGAGAC 974
QY 981 TGCACCTGATGAGGAGGCTCTCATAAAGTTCTGTGTAAGAAATATGTTTCAACGAAGA 1040
    |||
Db 975 TGCACCTGATGAGGAGGCTCTCATAAAGTTCTGTGTAAGAAATATGTTTCAACGAAGA 1034
QY 1041 TCGGGTGACAAAGGCCATAGAGAGATCAAAATTCGCCAAGATAAATTCGTGCCAAGAAAG 1100
    |||
Db 1035 TCGGGTGACAAAGGCCATAGAGAGATCAAAATTCGTGCCAAGATAAATTCGTGCCAAGAAAG 1094
QY 1101 ACTCGAGCTCTTTTCAACCAACTGCGCACATCAGACCGCTTAAACGGAGAGAGAC 1160
    |||
Db 1095 ACTCGAGCTCTTTTCAACCAACTGCGCACATCAGACCGCTTAAACGGAGAGAGAC 1154
QY 1161 TTCGGATTAACAGCAAGGCAAGCTGCGAACAAGAAAGAAAGGCTGTGGAAGAAGAA 1220
    |||
Db 1155 TTCGGATTAACAGCAAGGCAAGCTGCGAACAAGAAAGAAAGGCTGTGGAAGAAGAA 1214
QY 1221 AATATCTTGGATGCTTGAATGTACACTACGACTACGAAAGCAGCGGTGGCGTATGACTT 1280
    |||
Db 1215 AATATCTTGGATGCTTGAATGTACACTACGACTACGAAAGCAGCGGTGGCGTATGACTT 1274
QY 1281 CGCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTGGTAAAGTTGGTGCATGTT 1340
    |||
Db 1275 CGCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTGGTAAAGTTGGTGCATGTT 1334
QY 1341 TCAAGCTGGGGTAAGTTAGTTGTGTTGAAGATTTGTACCAAGTAACAAACTTAT 1400
    |||
Db 1335 TCAAGCTGGGGTAAGTTAGTTGTGTTGAAGATTTGTGTACCAAGTAACAAACTTAT 1394
QY 1401 GCCTGTTTTTACTCTCTGCTCTTGAAGTA 1431
    |||
Db 1395 GCCTGTTTTTACTCTCTGCTCTTGAAGTA 1425

RESULT 3
US-09-426-557-5
; Sequence 5, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 5
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(1173)
US-09-426-557-5

Query Match      86.6%; Score 1267; DB 4; Length 1381;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 GCCACAGCCCGCCGACAGCGATGGGCATCAAGGTTTGCAGAACTGCTGGCGGACAT 123
    |||
Db 16 GCCACAGCCCGCCGACAGCGATGGGCATCAAGGTTTGCAGAACTGCTGGCGGACAT 75
QY 124 GCGCCCAAGGCGATGAAGAGCAGAAAGTTGAGAGCTACTTCGCGCCCAAAATCGCGTGC 183
    |||
Db 76 GCGCCCAAGGCGATGAAGAGCAGAAAGTTGAGAGCTACTTCGCGCCCAAAATCGCGTGC 135
QY 184 GAGCCAGCATTGACATATACCATTCCTGATTGTAGTTGGAAGACAGCATGAAACT 243
    |||
Db 136 GAGCCAGCATTGACATATACCATTCCTGATTGTAGTTGGAAGACAGCATGAAACT 195
QY 244 CTCACAAATGAACCTGGTGAAGTCATGACATTTGCAAGGATGTTCAACCGGACAAATA 303
    |||
Db 196 CTCACAAATGAACCTGGTGAAGTCATGACATTTGCAAGGATGTTCAACCGGACAAATA 255
QY 304 AGATTACTGMAACGGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCCTCGTATATG 363
    |||
Db 256 AGATTACTGMAACGGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCCTCGTATATG 315
QY 364 AAGAAACAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATCAACCAAGATCTGACT 423
    |||
Db 316 AAGAAACAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATCAACCAAGATCTGACT 375
QY 424 GAGCGATGAGTAGAGATTAAGATGCGATTGAAAAATTTGCAAGAGAGCATGTAAG 483
    |||
Db 376 GAGCGATGAGTAGAGATTAAGATGCGATTGAAAAATTTGCAAGAGAGCATGTAAG 435
QY 484 GTCCACAGGCAACACACAGAGATTTGAACGCTATTAAGACTTTAGGGGTTCTCTGTT 543
    |||
Db 436 GTCCACAGGCAACACACAGAGATTTGAACGCTATTAAGACTTTAGGGGTTCTCTGTT 495
QY 544 GTAGAGGACCTTCTGAAGCAGACGAATGTGCAAGCCTTGTGATTAACGATTAAGT 603
    |||
Db 496 GTAGAGGACCTTCTGAAGCAGACGAATGTGCAAGCCTTGTGATTAACGATTAAGT 555
QY 604 TTCGCTGTTGCTTCAAGAGATATGACCTCCCTACTTTGGGGCTCCACGGTCCCTTGGT 663
    |||
Db 556 TTCGCTGTTGCTTCAAGAGATATGACCTCCCTACTTTGGGGCTCCACGGTCCCTTGGT 615
QY 664 CATTTAATGATCCAGTTCCAGAAATAACCTGTGATGGAATTTGATGTGGCAAGTT 723
    |||
Db 616 CATTTAATGATCCAGTTCCAGAAATAACCTGTGATGGAATTTGATGTGGCAAGTT 675
QY 724 TTGGAGAGCTTGAATCTCACATGACCAAGTTCATTTGATTTGTGCAATCCTGTGATGT 783
    |||
Db 676 TTGGAGAGCTTGAATCTCACATGACCAAGTTCATTTGATTTGTGCAATCCTGTGATGT 735
QY 784 GACTATTGTGATAGCATCAAGATATCGGGGGGCAACAGCTGTGAACCTTATTCGTCAA 843
    |||
Db 736 GACTATTGTGATAGCATCAAGATATCGGGGGGCAACAGCTGTGAACCTTATTCGTCAA 795
QY 844 CATGGTCCATAGAAAGCATCTTGGAGAACTCTAATTAAGACAGATATCAAAATTCCTGAG 903
    |||
Db 796 CATGGTCCATAGAAAGCATCTTGGAGAACTCTAATTAAGACAGATATCAAAATTCCTGAG 855
QY 904 GACTGGCCTTACCAAGAGCTGAGCCTGTTTCAAGAGACCTTAATGTACATTTGATATT 963
    |||
Db 856 GACTGGCCTTACCAAGAGCTGAGCCTGTTTCAAGAGACCTTAATGTACATTTGATATT 915
```

QY	964	CCCTAGCCTAAATATGAGACTGCACCTCATGAGAGAGGCTCTATTAAGTTTCTCGTATAAAGAT	1023
Db	916	CCTAGGCTAAAAATATGACTGTGCACCTCATGAGAGAGGCTCTATTAAGTTTCTCGTATAAAGAT	975
QY	1024	AATGTTTCAACGAAAGATGTGGGTGACAAAGGCCATGAGACATCAAAATCTCCAAAGAT	1083
Db	976	AATGTTTCAACGAAAGATGTGGGTGACAAAGGCCATGAGACATCAAAATCTCCAAAGAT	1035
QY	1084	AAATCTGCGCAAGGAAAGACTGCGAGTCCCTTTTCAAGCCAACTGCCACCATCAGCACCG	1143
Db	1036	AAATCTGCGCAAGGAAAGACTGCGAGTCCCTTTTCAAGCCAACTGCCACCATCAGCACCG	1095
QY	1144	CTAAACCGGAAGAGACTTGGGATTAACCAAGCAGACGGCTGCGCAACAAGAAAAAACAAG	1203
Db	1096	CTAAACCGGAAGAGACTTGGGATTAACCAAGCAGACGGCTGCGCAACAAGAAAAAACAAG	1155
QY	1204	GCTGCTGGAAGAAGAAATATCTTGGATCCTGATGTCAACTACGACTACGAAAGCAG	1263
Db	1156	GCTGCTGGAAGAAGAAATATCTTGGATCCTGATGTCAACTACGACTACGAAAGCAG	1215
QY	1264	CGGTGGCGTGATCACTTCGCTTAGATTAATTAACTCCCTGTTTAACTCAGAGCTTTGGT	1323
Db	1216	CGGTGGCGTGATCACTTCGCTTAGATTAATTAACTCCCTGTTTAACTCAGAGCTTTGGT	1275
QY	1324	AAAAGTT 1330	
Db	1276	AAAAGTT 1282	

```

? RESULT 4
? US-09-426-557-7
? Sequence 7, Application US/09426557
? Patent No. 6232527
? GENERAL INFORMATION:
? APPLICANT: Mahajan, Pramod B.
? TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and uses
? TITLE OF INVENTION: Thereof
? FILE REFERENCE: 0961
? CURRENT APPLICATION NUMBER: US/09/426,557
? CURRENT FILING DATE: 1999-10-22
? EARLIER APPLICATION NUMBER: 60/112,332
? EARLIER FILING DATE: 1998-12-15
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 7
? LENGTH: 1478
? TYPE: DNA
? ORGANISM: Zea mays
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (97)...(1233)
? US-09-426-557-7

```

Query Match	61.1%;	Score 894;	DB 4;	Length 1478;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 124;	Conservative	0;	Mismatches 7;	Indels 0;
				Gaps 0;

Qy	20	CGCGGTTTCTTGCGCATCCGGCTCAGGCGCGCGCGCGACGCCGACAGCGCGCGAG	79
Db	32	CGCGGTTTCTTGCGGCACCTCGGCTCAGCCCGCGCGCGACCCCGCAAGCCGCCGAG	91
Qy	80	ACGAGATGGCATCAAGGGTTTGAAGCAACCTGCTGGCGGCAATGCGCCCAAGGGCATGA	139
Db	92	ACGAGATGGCATCAAGGCTTTGACGAACTGCTGGCGGACATGGGCCCAAGGGCATGA	151
Qy	140	AGGAGCAGAAGTTGAGAGCTACTTGGGCGCAAAATCGCCGTGAGCGCCAGCATGACA	199
Db	152	AGGAGCAGAAGTTGAGAGCTACTTCGGCGCAAAATCGCCGTGAGCGCCAGCATGACA	211
Qy	200	TATACCACTTCTCTGATTTAGTTGGAAGGACAGGCATGAAACTCTCACAAAATGAAGCTG	259
Db	212	TCTACCACTTCTCTGATTTAGTTGGAAGGACAGGCATGAAACTCTCACAAAATGAAGCTG	271

Qy	260	GTAAAGTCACATAGCATTTTCGCAAGGAATGTCAACCGGACATTAAGTTATCTCGGAAGCGG	319
Db	272	GTGAAGTCATCAGTCATTTGCAAGGAATGTCAACCGGACAATAAGATTACTGGGAAGCGG	331
Qy	320	GAATCAAGCCAGTTTATGTTTGGATGGCAAGCCCTCTATATGAAAGAACAGAGCTTG	379
Db	332	GAATCAAGCCAGTTTATGTTTGGATGGCAAGCCCTCTATATGAAAGAACAGAGCTTG	391
Qy	380	CTAAAGTACTCAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCAGTACAGGTG	439
Db	392	CTAAAGTACTCTAAAAAGAGATGATGCAACCAAGATCTGACTGAGGCAGTACAGGTG	451
Qy	440	GAGATAAAGATGGCATTTGAAAAAATTGGACAGAGGAGCTTTAAAGGTCACAAAGCAACCA	499
Db	452	GAGATAAAGATGGCATTTGAAAAAATTGGACAGAGGAGCTTTAAAGGTCACAAAGCAACCA	511
Qy	500	ACGAAGATTGTTAAACGGCTTTTAAGACTTATGGGGGTTCTGTTGTAGAGGACCTTCTG	559
Db	512	ACGAAGATTGTTAAACGGCTTTTAAGACTTATGGGGGTTCTGTTGTAGAGGACCTTCTG	571
Qy	560	AAGCAGAAGCGAATGTGCGACCCCTTGGCTAATACGATAAGGTGTTCCGTGTGGCTTAG	619
Db	572	AAGCAGAAGCGAATGTGCGACCCCTTGGCTAATACGATAAGGTGTTCCGTGTGGCTTAG	631
Qy	620	AAGATATGATCCCTTACTACTTTTGGGGCTCCACAGTGTCCCTGTCATTAAATGATCCAA	679
Db	632	AAGATATGATCCCTTACTACTTTTGGGGCTCCACAGTGTCCCTGTCATTAAATGATCCAA	691
Qy	680	GTTCCAGAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTGGCAGGAGCTTGAAC	739

Qy	740	TCACCATGGACCACTTCATTGATTTGTGCACTCCTGTGTGGATGTGACTATTTGTGATTACA	799
Db	752	TCACCATTGACCACTGATTCATTGATTTGTGCACTCCTGTGTGGATGTGACTATTTGTGATTACA	811
Qy	800	TCGAAGGATGCGGGGGGCAAAACACCTGTGAACTTTATTTGTCAACATGGGTGCATATAGAA	859
Db	812	TCGAAGGATGCGGGGGGCAAAACACCTGTGAACTTTATTTGTCAACATGGGTGCATATAGAA	871
Qy	860	GCATTTGTGGAATCTTAATTAAGACGATATCAAAATTCCTGAGCACTGCGCTTACCAAG	919
Db	872	GCATTTGTGGAATCTTAATTAAGACGATATCAAAATTCCTGAGCACTGCGCTTACCAAG	931
Qy	920	AAGCTGACGCTTGTTCAGAGAGCCTATATGTCAATTTGGATTTCTGTAGCTTAAATATGA	979
Db	932	AAGCTGACGCTTGTTCAGAGAGCCTATATGTCAATTTGGATTTCTGTAGCTTAAATATGA	991
Qy	980	CTGCACCTGATGAGAGGGGTCTCAATAGTTTCTGGTAAAGATATATGGTTTCAACAGAG	1039
Db	992	CTGCACCTGATGAGAGGGGTCTCAATAGTTTCTGGTAAAGATATATGGTTTCAATGAG	1051
Qy	1040	ATCGGGGTACAAAGGCCATAGAGAGATCAAAATCTGGCAAGATTAATGTCGCAGAGAA	1099
Db	1052	ATCGGGGTACAAAGGCCATAGAGAGATCAAAATCTGGCAAGATTAATGTCGCAGAGAA	1111
Qy	1100	GACTGAGTCTCTTTTTCAGGCAACTCCACACATATGACACCGCTTAAACGGAAGAGA	1159
Db	1112	GACTGAGTCTCTTTTTCAGGCAACTCCACACATATGACACCGCTTAAACGGAAGAGA	1171
Qy	1160	CTTGGGATTAACAAGCAGGACGCTGCGAACAAGAAACAAAGGCTGGTGGAAAGAAGA	1219
Db	1172	CTTGGGATTAACAAGCAGGACGCTGCGAACAAGAAACAAAGGCTGGTGGAAAGAAGA	1231
Qy	1220	AATTAATCTTTGATGCTTGATGTATCACTACGACTAGCAAAAGCAGCGGTGGC	1270
Db	1232	AATTAATCTTTGATGCTTGATGTATCACTACGACTAGCAAAAGCAGCGGTGGC	1282

RESULT 5
US-08-757-046A-5
; Sequence 5, Application US/08757046A
; Patent No. 5876995

GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,046A
FILING DATE: 11-25-96
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/597,274
FILING DATE: 02-06-96
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6680-105B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 115...702
OTHER INFORMATION: apoaequorin-encoding gene
PUBLICATION INFORMATION:
AUTHORS: Imouye et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 82
PAGES: 3154-3158
DATE: (1985)
DOCUMENT NUMBER: PATENT NO.: 5,093,240
US-08-757-046A-5

Query Match 2.5%; Score 37; DB 2; Length 958;
Best Local Similarity 100.0%; Pred. No. 6,2e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 6
US-09-447-208-5
Sequence 5, Application US/09447208
Patent No. 6113886
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/447,208
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 0909/135,988
FILING DATE: 08-17-98
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/757,046
FILING DATE: 11-25-96
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/597,274
FILING DATE: 02-06-96
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24727-105C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-450-8499
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 115...702
OTHER INFORMATION: apoaequorin-encoding gene
PUBLICATION INFORMATION:
AUTHORS: Imouye et al.
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 82
PAGES: 3154-3158
DATE: (1985)
US-09-447-208-5

Query Match 2.5%; Score 37; DB 3; Length 958;
Best Local Similarity 100.0%; Pred. No. 6,2e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 7
US-09-135-988-5
Sequence 5, Application US/09135988
Patent No. 6152358
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce

```
? TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Heller Ehrman White & McNalliffe
? STREET: 4250 Executive Square, 7th Floor
? CITY: La Jolla
? STATE: CA
? COUNTRY: USA
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/135,988
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/757,046
? FILING DATE: 11-25-96
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/597,274
? FILING DATE: 02-06-96
? ATTORNEY/AGENT INFORMATION:
? NAME: Seidman, Stephanie L
? REGISTRATION NUMBER: 33,779
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-450-8400
? TELEFAX: 619-450-8499
? TELEX:
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 958 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE:
? ORIGINAL SOURCE:
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 115...702
? OTHER INFORMATION: apoaeguorin-encoding gene
? PUBLICATION INFORMATION:
? PATENT NO.: 5,093,240
? AUTHORS: Inouye et al.
? JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
? VOLUME: 82
? PAGES: 3154-3158
? DATE: (1985)
? US-09-135-988-5

Query Match 2.5%; Score 37; DB 3; Length 958;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||||||
DB 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927
```

```
RESULT 8
US-09-277-716-5
; Sequence 5, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
; ATTORNEY/AGENT INFORMATION:
; APPLICANT: Szent-Gyorgyi, Christopher
```

```
? APPLICANT: PROLUME, LTD.
? TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
? CURRENT APPLICATION NUMBER: US/09/277,716A
? EARLIER FILING DATE: 1998-03-26
? EARLIER APPLICATION NUMBER: 60/102,939
? EARLIER FILING DATE: 1998-10-01
? EARLIER APPLICATION NUMBER: 60/089,367
? EARLIER FILING DATE: 1998-06-15
? EARLIER APPLICATION NUMBER: 60/079,624
? EARLIER FILING DATE: 1998-03-27
? NUMBER OF SEQ ID NOS: 32
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 5
? LENGTH: 958
? TYPE: DNA
? ORGANISM: Aequorea (luminescent jellyfish)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (115)..(702)
? FEATURE:
? OTHER INFORMATION: apoaeguorin-encoding gene
? PUBLICATION INFORMATION:
? PATENT DOCUMENT NUMBER: 5,093,240
? PATENT FILING DATE: 1987-10-08
? PUBLICATION DATE: 1992-03-03
? PUBLICATION INFORMATION:
? AUTHORS: Inouye, S.
? TITLE: Cloning and sequence analysis of cDNA for the luminescent protein aequ
? JOURNAL: Proc. Natl. Acad. Sci. USA
? VOLUME: 82(10)
? PAGES: 3154-3158
? DATE: 1985-05
? US-09-277-716-5
```

```
Query Match 2.5%; Score 37; DB 4; Length 958;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||||||
DB 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927
```

```
RESULT 9
US-08-597-274A-5
; Sequence 5, Application US/08597274A
; Patent No. 6247995
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT NOVELTY ITEMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,274A
; FILING DATE: 02/06/96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
```

```

: REFERENCE/DOCKET NUMBER: 6680-105
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-238-0999
: TELEFAX: 619-238-0062
:
: TELEX:
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE:
: ORIGINAL SOURCE:
:
: FEATURE:
: NAME/KEY: Coding Sequence
: LOCATION: 115...702
: OTHER INFORMATION: apoaeguorin-encoding gene
: PUBLICATION INFORMATION:
: DOCUMENT NUMBER: 5,093,240
:
: AUTHORS: Inouye et al.
: JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
: VOLUME: 82
: PAGES: 3154-3158
: DATE: (1985)
:
: US-08-597-274A-5
:
Query Match      2.5%; Score 37; DB 4; Length 958;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
DB 891 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 927

RESULT 10
US-08-747-221B-24
: Sequence 24, Application US/08747221B
: Patent No. 6063610
: GENERAL INFORMATION:
: APPLICANT: Silver, Gary W.
: APPLICANT: Wisniewski, Nancy
: TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
: TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carol Talkington Verser, Ph.D.
: ADDRESSEE: Heska Corporation
: STREET: 1825 Sharp Point Drive
: CITY: Fort Collins
: STATE: Colorado
: COUNTRY: USA
: ZIP: 80525
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: Windows 95
: SOFTWARE: Wordperfect for Windows, Version 7.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/747,221B
: FILING DATE: No. 6063610ember 12, 1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Verser, Carol Talkington
: REGISTRATION NUMBER: 37,459
: REFERENCE/DOCKET NUMBER: FC-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 970/493-7272
: TELEFAX: 970/484-9505
: INFORMATION FOR SEQ ID NO: 24:
```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 2836 nucleotides
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 99..1889
:
: US-08-747-221B-24
:
Query Match      2.5%; Score 37; DB 3; Length 2836;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
DB 2781 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2817

RESULT 11
US-08-747-221B-26/c
: Sequence 26, Application US/08747221B
: Patent No. 6063610
: GENERAL INFORMATION:
: APPLICANT: Silver, Gary W.
: APPLICANT: Wisniewski, Nancy
: TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
: TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carol Talkington Verser, Ph.D.
: ADDRESSEE: Heska Corporation
: STREET: 1825 Sharp Point Drive
: CITY: Fort Collins
: STATE: Colorado
: COUNTRY: USA
: ZIP: 80525
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: Windows 95
: SOFTWARE: Wordperfect for Windows, Version 7.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/747,221B
: FILING DATE: No. 6063610ember 12, 1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Verser, Carol Talkington
: REGISTRATION NUMBER: 37,459
: REFERENCE/DOCKET NUMBER: FC-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 970/493-7272
: TELEFAX: 970/484-9505
: INFORMATION FOR SEQ ID NO: 26:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2836 nucleotides
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
:
: US-08-747-221B-26
:
Query Match      2.5%; Score 37; DB 3; Length 2836;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
      |||
DB 56 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20

RESULT 12
```

US-09-005-051-24
; Sequence 24, Application US/09005051
; Patent No. 6291222
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,051
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/747,221
; FILING DATE: No. 6291222el December 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2836 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 99..1889
; US-09-005-051-24
Query Match 2.5%; Score 37; DB 4; Length 2836;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 2781 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2817
RESULT 13
US-09-005-051-26/C
; Sequence 26, Application US/09005051
; Patent No. 6291222
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado

COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,051
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/747,221
; FILING DATE: No. 6291222el December 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2836 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-005-051-26
Query Match 2.5%; Score 37; DB 4; Length 2836;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1427 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 56 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20
RESULT 14
US-08-487-135B-1
; Sequence 1, Application US/08487135B
; Patent No. 5821122
; GENERAL INFORMATION:
; APPLICANT: Yannick Gulloux; Francine Jotereau;
; APPLICANT: Thierry Boon-Falleur; Sophie Lucas;
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES PEPTIDES
; TITLE OF INVENTION: WHICH FORM COMPLEXES WITH MHC MOLECULE HLA-A2
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,135B
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,135
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5821122man D
; REGISTRATION NUMBER: 30,946


```

; REFERENCE/DOCKET NUMBER: LUD 5388
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-487-135B-1

Query Match 2.5%; Score 36; DB 1; Length 2237;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2196 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2231

RESULT 15
US-08-915-972A-1
; Sequence 1, Application US/08915972A
; Patent No. 5886145
; GENERAL INFORMATION:
; APPLICANT: Yannick Gulloux; Francine Jotereau;
; APPLICANT: Thierry Boon-Falleur; Sophie Lucas;
; APPLICANT: Vincent Brichard
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
; TITLE OF INVENTION: COMPLEXES WITH MHC MOLECULE HLA-A2 AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,972A
; FILING DATE: August 21, 1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,135
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5886145man D
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5388 - JEL/NDH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-915-972A-1

Query Match 2.5%; Score 36; DB 2; Length 2237;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2196 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2231
```

```

RESULT 16
US-09-177-909-1
; Sequence 1, Application US/09177909
; Patent No. 5958711
; GENERAL INFORMATION:
; APPLICANT: Yannick Gulloux; Francine Jotereau;
; APPLICANT: Thierry Boon-Falleur; Sophie Lucas;
; APPLICANT: Vincent Brichard
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES PEPTIDES WHICH
; TITLE OF INVENTION: FORM COMPLEXES WITH MHC MOLECULE HLA-A2 AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/177,909
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/915,972
; FILING DATE: August 21, 1997
; APPLICATION NUMBER: 08/487,135
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5958711man D
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5388 - JEL/NDH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2237
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-177-909-1

Query Match 2.5%; Score 36; DB 2; Length 2237;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 2196 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2231

RESULT 17
US-08-171-385-27/C
; Sequence 27, Application US/0811385
; Patent No. 5527884
; GENERAL INFORMATION:
; APPLICANT: Mary E. Russell
; APPLICANT: Ulrike Utans
; TITLE OF INVENTION: Mediators of Chronic Allograft
; TITLE OF INVENTION: Rejection
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
```

COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,385
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/006001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 340
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-171-385-27

Query Match 2.4%; Score 35; DB 1; Length 340;
Best Local Similarity 100.0%; Pred. No. 4,8e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 39 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 18
US-08-361-441B-27/c
Sequence 27, Application US/08361441B
Patent No. 6077948
GENERAL INFORMATION:
APPLICANT: Russell, Mary E.
APPLICANT: Utans, Ulrike
TITLE OF INVENTION: MEDIATORS OF CHRONIC ALLOGRAFT REJECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/361,441B
FILING DATE: 21-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/171,385
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906

TELEX: 200154
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-361-441B-27

Query Match 2.4%; Score 35; DB 3; Length 340;
Best Local Similarity 100.0%; Pred. No. 4,8e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 39 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 19
US-08-104-072B-2
Sequence 2, Application US/08104072B
Patent No. 5639948
GENERAL INFORMATION:
APPLICANT: Michiels, Frank
APPLICANT: Morioka, Sinji
APPLICANT: Scheirlinck, Trees
APPLICANT: Komari, Toshiko
TITLE OF INVENTION: Stamen-specific Promoters from Rice
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 3100 No. 563948west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/104,072B
FILING DATE: 05-AUG-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 9200272
FILING DATE: 06-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91403352.7
FILING DATE: 10-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91402590.3
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 91400318.1
FILING DATE: 08-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kowalczyk, Katherine M.
REGISTRATION NUMBER: 36,848
REFERENCE/DOCKET NUMBER: 8076.93USWO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: rice

;; SEQ ID NO 262
;; LENGTH: 593
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc-feature
;; LOCATION: (1)...(593)
;; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-262

Query Match 2.4%; Score 35; DB 4; Length 593;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 74 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 40

RESULT 23
US-08-713-000-8

;; Sequence 8, Application US/08713000
;; Patent No. 5850020
;; GENERAL INFORMATION:
;; APPLICANT: Bloksberg, Leonard N.
;; APPLICANT: Havukkala, Ilkka
;; APPLICANT: Grierson, Alastair
;; TITLE OF INVENTION: MATERIALS AND METHODS FOR THE
;; TITLE OF INVENTION: MODIFICATION OF PLANT LIGNIN CONTENT
;; NUMBER OF SEQUENCES: 15
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: Spekman Picard PLLC
;; STREET: 2601 Elliott Avenue, Suite 4185
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98121
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/713,000
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sleath, Janet
;; REGISTRATION NUMBER: 37,007
;; REFERENCE/DOCKET NUMBER: 11000.1003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-269-0565
;; TELEFAX: 206-269-0563
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 740 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-713-000-8

Query Match 2.4%; Score 35; DB 2; Length 740;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 690 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724

RESULT 24
US-08-975-316-8
;; Sequence 8, Application US/08975316

;; Patent No. 5952486
;; GENERAL INFORMATION:
;; APPLICANT: BLOKSBERG, Leonard N., HAVUKKALA, Ilkka
;; APPLICANT: and GRIERSON, Alastair W.
;; TITLE OF INVENTION: MATERIALS AND METHODS FOR
;; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
;; NUMBER OF SEQUENCES: 88
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: Law Offices of Ann W. Spekman
;; STREET: 2601 Elliott Avenue, Suite 4185
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98121
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/975,316
;; FILING DATE:
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/713,000
;; FILING DATE: September 11, 1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SLEATH, Janet
;; REGISTRATION NUMBER: 37,007
;; REFERENCE/DOCKET NUMBER: 11000/1003C1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-269-0565
;; TELEFAX: 206-269-0563
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 740 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-975-316-8

Query Match 2.4%; Score 35; DB 2; Length 740;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 690 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724

RESULT 25
US-09-211-710-8
;; Sequence 8, Application US/09211710A
;; Patent No. 6204434
;; GENERAL INFORMATION:
;; APPLICANT: Bloksberg, Leonard N.
;; APPLICANT: Havukkala, Ilkka
;; APPLICANT: Grierson, Alastair
;; TITLE OF INVENTION: Materials and Methods for the
;; TITLE OF INVENTION: Modification of Plant Lignin Content
;; FILE REFERENCE: 11000.1003c3
;; CURRENT APPLICATION NUMBER: US/09/211,710A
;; CURRENT FILING DATE: 1998-12-14
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8
;; LENGTH: 740
;; TYPE: DNA
;; ORGANISM: Pinus radiata
US-09-211-710-8

Query Match 2.4%; Score 35; DB 4; Length 740;

Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 690 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724

RESULT 26
US-08-975-316-58
; Sequence 58, Application US/08975316
; Patent No. 5952486
; GENERAL INFORMATION:
; APPLICANT: BLOKSBERG, Leonard N., HAYUKKALA, Ilkka
; APPLICANT: and GRIERSON, Alastair W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975.316
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,000
; FILING DATE: September 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: SLEATH, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000/1003C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 741 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-975-316-58
Query Match 2.4%; Score 35; DB 2; Length 741;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 690 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724
RESULT 27
US-08-400-006B-6
; Sequence 6, Application US/08400006B
; Patent No. 6229065
; GENERAL INFORMATION:
; APPLICANT: FREYSSINET, Georges
; APPLICANT: SATILAND, Alain
; TITLE OF INVENTION: PRODUCTION OF PLANTS RESISTANT TO
; ATTACKS BY SCLEROTINIA SCLEROTIIFORM BY THE INTRODUCTION OF
; A GENE ENCODING AN OXALATE OXIDASE
; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 1737 King Street, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22314-2756
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,006B
; FILING DATE: 06-Mar-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91/02874
; FILING DATE: 05-MAR-1991
; APPLICATION NUMBER: PCT/FR92/00195
; FILING DATE: 04-MAR-1992
; APPLICATION NUMBER: US 07/941,135
; FILING DATE: 03-DEC-1992
; APPLICATION NUMBER: 08/207,105
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Teresa Stanek Rea
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 022650-189
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1075
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-400-006B-6
Query Match 2.4%; Score 35; DB 4; Length 1075;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 1011 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1045
RESULT 28
US-08-991-789A-169
; Sequence 169, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,789A
FILING DATE: 11-Dec-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 210121.419C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 1265 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-991-789A-169
SEQUENCE DESCRIPTION: SEQ ID NO: 169:
Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 1265;
Pred. No. 4.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1226 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260
RESULT 29
US-09-062-451-169
Sequence 169, Application US/09062451
Patent No. 6344550
GENERAL INFORMATION:
APPLICANT: Frudakis, Tony N.
APPLICANT: Smith, John M.
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 297
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,451
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 1265 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
```

```
TOPOLOGY: linear
US-09-062-451-169
Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 1265;
Pred. No. 4.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1226 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1260
RESULT 30
US-09-721-822A-10
Sequence 10, Application US/09721822A
Patent No. 6306606
GENERAL INFORMATION:
APPLICANT: Michael J. Weber
APPLICANT: Jacqueline Wyatt
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF MP-1 EXPRESSION
FILE REFERENCE: RUS-0142
CURRENT APPLICATION NUMBER: US/09/721,822A
CURRENT FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 135
SEQ ID NO 10
LENGTH: 1315
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (147)...(521)
OTHER INFORMATION:
US-09-721-822A-10
Query Match
Best Local Similarity 100.0%; Score 35; DB 4; Length 1315;
Pred. No. 4.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
Db 1261 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1295
RESULT 31
US-09-262-749-1
Sequence 1, Application US/09262749
Patent No. 6261793
GENERAL INFORMATION:
APPLICANT: Whyte, David
APPLICANT: McGuirk, Marjole
APPLICANT: Nunez-Oliva, Irma
APPLICANT: Hockenderry, Tish
APPLICANT: Pal, James
TITLE OF INVENTION: RAS CONVERTING ENDOPROTEASE (RCE) AND
METHODS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road K-6-1-1990
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: MACOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/262,749
FILING DATE:
CLASSIFICATION:
```

```

? ATTORNEY/AGENT INFORMATION:
?   NAME: /Thampee, Immac J
?   REGISTRATION NUMBER: 36,322
?   REFERENCE/DOCKET NUMBER: OC0100505
?   TELECOMMUNICATION INFORMATION:
?   TELEPHONE: (908) 298-5061
?   TELEFAX: (908) 298-5388
?   INFORMATION FOR SEQ ID NO: 1:
?     SEQUENCE CHARACTERISTICS:
?       LENGTH: 1483 base pairs
?       TYPE: nucleic acid
?       STRANDEDNESS: single
?       TOPOLOGY: linear
?       MOLECULE TYPE: cDNA
?       HYPOTHETICAL: NO
?       ANTI-SENSE: NO
?       ORIGINAL SOURCE:
?         ORGANISM: Homo sapiens
?         FEATURE:
?           NAME/KEY: CDS
?           LOCATION: 14..1000
?   OS-09-262-749-1

```

Query Match	2.4%	Score 35;	DB 4;	Length 1483;
Best Local Similarity	100.0%;	Pred. No. 4.3e-06;		
Matches 35; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	1429	GTAAATAAAAAAAAAAAAAAAAAAAAAA	1463
		T	
Db	1432	GTAAATAAAAAAAAAAAAAAAAAAAAAA	1466

```

RESULT 32
US-09-413-574-1
Sequence 1, Application US/09413574
Patent No. 6235972
GENERAL INFORMATION:
APPLICANT: Manajan, Pramod B.
APPLICANT: Tagiliani, Laura
TITLE OF INVENTION: Make Rad23 Genes and Uses Thereof
FILE REFERENCE: 0964
CURRENT APPLICATION NUMBER: US/09/413,574
CURRENT FILING DATE: 1999-10-06
EARLIER APPLICATION NUMBER: 60/109,728
EARLIER FILING DATE: 1998-11-23
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1522
TYPE: DNA
ORGANISM: zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (38)...(1272)
US-09-413-574-1

```

Query Match	2.4%	Score 35;	DB 4;	Length 1522;
Best Local Similarity	100.0%	Pred. No. 4.3e-06;		
Matches 35; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1429	GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1463
Db	1458	GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1492

RESULT33
US-09-461-474-7
; Sequence 7, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime

```

1  TITLE OF INVENTION: Plant Metal Transporters
2  FILE REFERENCE: B91303 US NA
3  CURRENT APPLICATION NUMBER: US/09/461,474
4  CURRENT FILING DATE: 1999-12-14
5  EARLIER APPLICATION NUMBER: 60/112,562
6  EARLIER FILING DATE: 1998-12-16
7  NUMBER OF SEQ ID NOS: 17
8  SOFTWARE: Microsoft Office 97
9  SEQ ID NO 7
10 LENGTH: 1551
11 TYPE: DNA
12 ORGANISM: Zea mays
13 US-09-461-474-7

```

Query Match	2.48;	Score 35;	DB 4;	Length 1551;
Best Local Similarity	100.0%;	Pred. No. 4.2e-06;		
Matches 35; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy	1429	GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1463
Db	1517	GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1551

```

US-09-022-669-1
US-09-022-669-1
;; Sequence 1, Application US/09022669
;; Patent No. 61107717
;; GENERAL INFORMATION:
;; APPLICANT: KILLY, KRISTINE
;; APPLICANT: SOUTHAN, CHRISTOPHER
;; APPLICANT: KNAB, ANNE
;; TITLE OF INVENTION: Human RCE1
;; NUMBER OF SEQUENCES: 3
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: RATNER & PRESTIA
;; STREET: P.O. BOX 980
;; CITY: VALLEY FORGE
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19482
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FASTSO for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/022,669
;; FILING DATE: 12-FEB-1998
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 97304437.3
;; FILING DATE: JUNE 24, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: PRESTIA, PAUL F
;; REGISTRATION NUMBER: 23,031
;; REFERENCE/DOCKET NUMBER: GH-70379
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-407-0700
;; TELEFAX: 610-407-0701
;; TELEX: 846169
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1553 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
US-09-022-669-1

```

Query Match	2.4%;	Score 35;	DB 3;	Length 1553;
Best Local Similarity	100.0%;	Pred. No. 4.2e-06;		
Matches	35;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

Qy	1429 GTAAAAAAAAAAAAAAAAAAAAA	1463
Db	1516 GTAAAAAAAAAAAAAAAAAAAAA	1550

Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 1618 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1652

RESULT 38

US-08-463-694-7
; Sequence 7, Application US/08463694
; Patent No. 5696233

; GENERAL INFORMATION:

; APPLICANT: EVANS Ph.D., RONALD M.
; APPLICANT: MANGELSDORF Ph.D., DAVID J.
; APPLICANT: ONG Ms., ESTELITA S.
; APPLICANT: ORO Ph.D., ANTHONY E.
; APPLICANT: BORGMAYER Ph.D., UWE K.
; APPLICANT: GIGUERE Ph.D., VINCENT NMN
; APPLICANT: YAO Mr., TSO-PANG NMN
; TITLE OF INVENTION: NOVEL RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: US
; ZIP: 90071-2921

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,694
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/761,068
; FILING DATE: 17-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter Ph.D., Stephen E.

REGISTRATION NUMBER: 31192

; REFERENCE/DOCKET NUMBER: P31 8936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949

INFORMATION FOR SEQ. ID NO: 7:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: cDNA

; IMMEDIATE SOURCE:
; CLONE: XR2 (XR2.SEG)

FEATURE:

; NAME/KEY: CDS
; LOCATION: 148..1470

US-08-463-694-7

Query Match 2.4%; Score 35; DB 1; Length 1659;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 1618 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1652

RESULT 39
US-08-694-501-7

; Sequence 7, Application US/08694501
; Patent No. 5710004

; GENERAL INFORMATION:

; APPLICANT: EVANS Ph.D., RONALD M.
; APPLICANT: MANGELSDORF Ph.D., DAVID J.
; APPLICANT: ONG Ms., ESTELITA S.
; APPLICANT: ORO Ph.D., ANTHONY E.
; APPLICANT: BORGMAYER Ph.D., UWE K.
; APPLICANT: GIGUERE Ph.D., VINCENT NMN
; APPLICANT: YAO Mr., TSO-PANG NMN
; TITLE OF INVENTION: NOVEL RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: US
; ZIP: 90071-2921

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,501
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/333,358
; FILING DATE:
; APPLICATION NUMBER: US/07/761,068
; FILING DATE:

ATTORNEY/AGENT INFORMATION:

; NAME: Reiter Ph.D., Stephen E.
; REGISTRATION NUMBER: 31192
; REFERENCE/DOCKET NUMBER: P31 8936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949

INFORMATION FOR SEQ. ID NO: 7:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: cDNA

; IMMEDIATE SOURCE:
; CLONE: XR2 (XR2.SEG)

FEATURE:

; NAME/KEY: CDS
; LOCATION: 148..1470

US-08-694-501-7

Query Match 2.4%; Score 35; DB 1; Length 1659;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
DB 1618 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1652

RESULT 40

US-08-985-950-7
; Sequence 7, Application US/08985950

; GENERAL INFORMATION:

; APPLICANT: Adema, Gosse Jan
; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute

```

? STREET: 901 California Avenue
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1104
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/985,950
? FILING DATE: 05-DEC-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/041,279
? FILING DATE: 21-MARCH-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/033,181
? FILING DATE: 16-DEC-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/032,252
? FILING DATE: 06-DEC-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Chung, Edwin P.
? REGISTRATION NUMBER: 34,090
? REFERENCE/DOCKET NUMBER: DX0670K
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650)852-9196
? TELEFAX: (650)496-1204
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1728 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 69..929
? NAME/KEY: mat_peptide
? LOCATION: 132..929
?
? US-08-985-950-7
?
? Query Match 2.4%; Score 35; DB 3; Length 1728;
? Best Local Similarity 100.0%; Pred. No. 4.2e-06;
? Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
?
? QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
? Db 1674 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1708
?
? RESULT 41
? US-08-655-352-10
? Sequence 10, Application US/08655352
? Patent No. 6077991
? GENERAL INFORMATION:
? APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
? PRODUCTION OF MALE-STERILE PLANTS
? NUMBER OF SEQUENCES: 19
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Klargust Sparkman Campbell Leigh &
? ADDRESSEE: Whinston, LLP
? STREET: One World Trade Center
? STREET: 121 S.W. Salmon Street
? STREET: Suite 1600
? CITY: Portland
? STATE: Oregon
? COUNTRY: United States of America

```

```

? ZIP: 97204
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Disk, 3-1/2 inch
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: MS DOS
? SOFTWARE: Wordperfect 5.1
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/655,352
? FILING DATE:
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/323,449
? FILING DATE: October 14, 1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Dow, Alan, E.
? REGISTRATION NUMBER: 35,123
? REFERENCE/DOCKET NUMBER: 4630-45000
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (503) 226-7391
? TELEFAX: (503) 228-9446
? INFORMATION FOR SEQ ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1776 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double stranded
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA to mRNA
? DESCRIPTION: Tobacco Ccank cDNA and deduced amino-acid
? FEATURE:
? NAME/KEY: protein-coding sequence (not including
? stop codon)
? NAME/KEY: stop codon
? LOCATION: nucleotides 20-1570
?
? US-08-655-352-10
?
? Query Match 2.4%; Score 35; DB 3; Length 1776;
? Best Local Similarity 100.0%; Pred. No. 4.2e-06;
? Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
?
? QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
? Db 1720 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1754
?
? RESULT 42
? US-08-606-505B-1
? Sequence 1, Application US/08606505B
? Patent No. 6114601
? GENERAL INFORMATION:
? APPLICANT: KIKUCHI, Yasuhiro
? APPLICANT: KIKUCHI, Shigeto
? APPLICANT: KIKUCHI, Shigeto
? APPLICANT: SHIMADA, Yukihisa
? APPLICANT: OHBAYASHI, Masaya
? APPLICANT: SHIMADA, Ritsuko
? APPLICANT: OKINAKA, Yasushi
? TITLE OF INVENTION: NOVEL PLANT GENES
? NUMBER OF SEQUENCES: 67
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FITZPATRICK, CELIA, HARPER & SCINTO
? STREET: 30 Rockefeller Plaza
? CITY: New York
? STATE: New York
? COUNTRY: U.S.A.
? ZIP: 10112-3801
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
? COMPUTER: IBM PS/4
? OPERATING SYSTEM: MS-DOS Ver3.30
? SOFTWARE: PATENT AID Ver1.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/606,505B
? FILING DATE: 23-MAR-1996
? PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, LAWRENCE S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 1 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1824 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Petunia hybrida
; STRAIN: Falcon Blue
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116 to 1633
; IDENTIFICATION METHOD: by experiment
;
US-08-606-505B-1

Query Match          2.4%; Score 35; DB 3; Length 1824;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAA... 1463
DB 1785 GTAAAAA... 1819

RESULT 43
US-09-616-990-1
; Sequence 1, Application US/09616990
; Patent No. 6232109
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; KIYOKAWA, Shigeto
; SHIMADA, Yukihisa
; OHBAYASHI, Masaya
; SHIMADA, Ritsuko
; OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS Ver3.30
; SOFTWARE: PATENT AID Ver1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/616,990
; FILING DATE: 14-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, LAWRENCE S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 1 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1824 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Petunia hybrida
; STRAIN: Falcon Blue
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116 to 1633
; IDENTIFICATION METHOD: by experiment
; SEQUENCE DESCRIPTION: SEQ ID NO: 1
;
US-09-616-990-1

Query Match          2.4%; Score 35; DB 4; Length 1824;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAA... 1463
DB 1785 GTAAAAA... 1819

RESULT 44
US-09-385-801-1
; Sequence 1, Application US/09385801
; Patent No. 6180850
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Shi, Jinrui
; TITLE OF INVENTION: Maize Ku70 Orthologue and Uses Thereof
; FILE REFERENCE: 0932
; CURRENT APPLICATION NUMBER: US/09/385,801
; EARLIER APPLICATION NUMBER: 60/098,986
; EARLIER FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2399
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)...(2181)
;
US-09-385-801-1

Query Match          2.4%; Score 35; DB 4; Length 2399;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAA... 1463
DB 2356 GTAAAAA... 2390

RESULT 45
US-09-234-332-1
; Sequence 1, Application US/09234332A
; Patent No. 6087168
; GENERAL INFORMATION:
; APPLICANT: Cedars-Sinai Medical Center
; APPLICANT: Michel F. Levesque, M.D.
; APPLICANT: Thomas Neuman, Ph.D.
; TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
; FILE REFERENCE: P07 41494
; CURRENT APPLICATION NUMBER: US/09/234,332A
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2502
```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: gene
; LOCATION: (0)...(0)
; OTHER INFORMATION: Neuro D1 gene: Genbank accession D62347
US-09-234-332-1

Query Match 2.4%; Score 35; DB 3; Length 2502;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 1822 GTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1856

Search completed: November 5, 2002, 13:48:45
Job time : 111.881 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 13:04:56 ; Search time 1251.78 Seconds
(without alignments)
15774.429 Million cell updates/sec

Title: US-09-805-311-1

Perfect score: 1463
Sequence: 1 caccgaataagctcgcgcc.....aaaaaaaaaaaaaaaa 1463

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrti:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	485	33.2	554	10 BE639421	BE639421 946033A02
C 2	378	25.8	550	10 BE639422	BE639422 946033A02
C 3	350	23.9	586	9 AI881599	AI881599 606068G09
C 4	334	22.8	467	9 BE186786	BE186786 946012C08
C 5	317	21.7	553	9 AM562789	AM562789 660065H06
C 6	295	20.2	470	9 AI861468	AI861468 614014D03
C 7	265	18.1	456	9 AI065689	AI065689 ag91f12.x
C 8	263	18.0	901	10 BC837708	BC837708 Zml0.01A0
C 9	249	17.0	474	9 AM559173	AM559173 660065H06
C 10	245	16.7	532	9 AM000375	AM000375 614014D03
C 11	241	16.5	475	9 AM562517	AM562517 660065H06
C 12	225	15.4	376	9 AI065546	AI065546 ag88e02.x
C 13	224	15.3	470	9 AM288831	AM288831 707010F11
C 14	192	13.1	414	9 AM288784	AM288784 707010C02
C 15	187	12.8	235	9 AM562788	AM562788 660065H06
C 16	148	10.1	470	9 AM288831	AM288831 707010F11
C 17	130	8.9	363	9 AM562518	AM562518 660065H06

C 18	103	7.0	232	10 BF727781	BF727781 1000052F1
C 19	93	6.4	553	9 AI834484	AI834484 606068G09
C 20	68	4.6	126	9 AM147048	AM147048 707010C02
C 21	65	4.4	126	9 AM147048	AM147048 707010C02
C 22	57	3.9	225	9 AI947478	AI947478 614047B01
C 23	41	2.8	346	10 BG406288	BG406288 sac25h02.
C 24	41	2.8	468	9 AL121294	AL121294 DKRP762L
C 25	41	2.8	642	10 BI761155	BI761155 603043646
C 26	41	2.8	822	10 BI760815	BI760815 603043846
C 27	40	2.7	202	9 AM307211	AM307211 sf54b11.y
C 28	40	2.7	203	9 AI373782	AI373782 q254906.x
C 29	40	2.7	294	9 AA678858	AA678858 ah05a03.s
C 30	40	2.7	409	10 BG301426	BG301426 kt04b11.y
C 31	40	2.7	411	10 BM378668	BM378668 MEST567-E
C 32	40	2.7	522	9 AA901800	AA901800 NCC1A3T7
C 33	40	2.7	741	10 BF134331	BF134331 601784314
C 34	39	2.7	109	9 AM568974	AM568974 s173909.y
C 35	39	2.7	150	10 BI319195	BI319195 949026H11
C 36	39	2.7	187	10 BF749502	BF749502 IL0-BN042
C 37	39	2.7	201	10 BF749503	BF749503 IL0-BN042
C 38	39	2.7	226	10 BF764516	BF764516 RCI-CS006
C 39	39	2.7	320	9 AM234203	AM234203 sf22d11.y
C 40	39	2.7	365	12 AZ953545	AZ953545 2M0218A19
C 41	39	2.7	431	9 AA121171	AA121171 z188905.s
C 42	39	2.7	432	9 AA121150	AA121150 z188d07.s
C 43	39	2.7	444	10 BM357517	BM357517 N16IV-E8
C 44	39	2.7	448	9 AI362270	AI362270 qy50e11.x
C 45	39	2.7	593	9 AL501532	AL501532 AL501532

ALIGNMENTS

RESULT 1
BE639421/c 554 bp mRNA linear EST 30-AUG-2000
LOCUS BE639421 946033A02.x2 946 - tassal primordium prepared by Schmidt lab zea
DEFINITION mays CDNA, mRNA sequence.
ACCESSION BE639421
VERSION BE639421.1 GI:9952838
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade: Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 554)
REFERENCE Walbot,V.
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.
FEATURES
source Location/Qualifiers
1..554
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/c1one_lib="946 - tassal primordium prepared by Schmidt lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOIR"
/note="Organ: tassels; Vector: HybridAP; Site_1: EcoRI; Site_2: XhoI; George Chuck dissected immature tassels between imm and 3mm. Sharon Stanfield prepared the cDNA

library in HybriZAP. Sample insert size range was 350 bp
to 3 kb with a 1 kb average."
BASE COUNT 145 a 131 c 110 g 168 t
ORIGIN

Query Match 33.2%; Score 485; DB 10; Length 554;
Best Local Similarity 99.8%; Pred. No. 8.6e-100;
Matches 535; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 896 TTCCTGAGACCTGGCTTACCAGAAAGCTGACGCTTTGTTCAAGAGCCTAATGTCACAT 955
|||||
Db 554 TTCCTGAGACCTGGCTTACCAGAAAGCTGACGCTTTGTTCAAGAGCCTAATGTCACAT 495
QY 956 TGGATATCTCGAGCTAAATGAGTGCACCTGATGAGGAGGCTCATAGTTTCTCTG 1015
|||||
Db 494 TGGATATCTCGAGCTAAATGAGTGCACCTGATGAGGAGGCTCATAGTTTCTCTG 435
QY 1016 TAAAGATTAATGGTTTCACGAAAGATCGGGTGACAAAGGCCATAGAGAAATCAATCTG 1075
|||||
Db 434 TAAAGATTAATGGTTTCACGAAAGATCGGGTGACAAAGGCCATAGAGAAATCAATCTG 375
QY 1076 CCAAGATTAATGTCGCAAGAGAGCTGAGTCTTTTCAAGCCAACTGCCACACAT 1135
|||||
Db 374 CCAAGATTAATGTCGCAAGAGAGCTGAGTCTTTTCAAGCCAACTGCCACACAT 315
QY 1136 CAGCACCGCTAAAGGAGAGAGCTTCGATTAACAAAGCAGCTGCCACACAA 1195
|||||
Db 314 CAGCACCGCTAAAGGAGAGAGCTTCGATTAACAAAGCAGCTGCCACACAA 255
QY 1196 AACCAAAAGCTGTGTGAAGAGAAATATCTTGATGCTGTGATGACACTACGACTAC 1255
|||||
Db 254 AACCAAAAGCTGTGTGAAGAGAAATATCTTGATGCTGTGATGACACTACGACTAC 195
QY 1256 GAAAGCAGCGGTGGCGTGATGACTTGGCTTAGATTTAATACCTCTGTTTAACTCA 1315
|||||
Db 194 GAAAGCAGCGGTGGCGTGATGACTTGGCTTAGATTTAATACCTCTGTTTAACTCA 135
QY 1316 GCTTTGTTAAAGTTTGCTATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGAT 1375
|||||
Db 134 GCTTTGTTAAAGTTTGCTATGTTTCAAGCTGGGGTAAGTTAGTTGTTTGAAGAGAT 75
QY 1376 TGGTGTACCAAGTACAAACTATCGCTGTTTCTTACTCTTGTCTTGAAGTA 1431
|||||
Db 74 TGGTGTACCAAGTACAAACTATCGCTGTTTCTTACTCTTGTCTTGAAGTA 19

RESULT 2
BE639422 550 bp mRNA linear EST 30-AUG-2000
LOCUS 946033A02.y2 946 - tassal primordium prepared by Schmidt lab Zea
DEFINITION
ACCESSION BE639422
VERSION BE639422.1 GI:9952839
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 550)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Contact: Walbot V
COMMENT Unpublished (1999)
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.
FEATURES
Location/Qualifiers

source 1. .550
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_id="946 - tassal primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="X10LR"
/note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 kb with a 1 kb average."
BASE COUNT 177 a 116 c 149 g 108 t
ORIGIN

Query Match 25.8%; Score 378; DB 10; Length 550;
Best Local Similarity 99.6%; Pred. No. 7.3e-76;
Matches 548; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 46 AGCCGCGCCGCCACCCGCCACAGCCGCGCAGACGATGGCATCAAGGTTGACG 105
|||||
Db 1 AGCCGCGCCGCCACCCGCCACAGCCGCGCAGACGATGGCATCAAGGTTGACG 60
QY 106 AACTGTGTGGGGCAATAGCGCCCAAGCGCATGAAAGAGCAGATCGAGCTACTTC 165
|||||
Db 61 AACTGTGTGGGGCAATAGCGCCCAAGCGCATGAAAGAGCAGATCGAGCTACTTC 120
QY 166 GCGCCAAATCGCGCGTGCAGCGCATGAGCATATACGATTCCTGATTGTAGTTGGA 225
|||||
Db 121 GCGCCAAATCGCGCGTGCAGCGCATGAGCATATACGATTCCTGATTGTAGTTGGA 180
QY 226 AGGACAGG-CATGGAACCTTCACAAATGAAGCTGTGAAGTCACTAGTCATTTGCAAG 284
|||||
Db 181 AGGACAGG-CATGGAACCTTCACAAATGAAGCTGTGAAGTCACTAGTCATTTGCAAG 240
QY 285 AATGTTCAACCGGCAATATAGATTACGAGCGGGAATCAAGCCATTATGTTTGA 344
|||||
Db 241 AATGTTCAACCGGCAATATAGATTACGAGCGGGAATCAAGCCATTATGTTTGA 300
QY 345 TGGCAGGCTCTCTGATATGAGAAACAAGAGCTTGTAAAGATATCAAAAAGAGATGA 404
|||||
Db 301 TGGCAGGCTCTCTGATATGAGAAACAAGAGCTTGTAAAGATATCAAAAAGAGATGA 360
QY 405 TGCACCAAGAATCTGACTGAGCAGTAGAGGTAGAGATTAAGATCCGATTGAAAATT 464
|||||
Db 361 TGCACCAAGAATCTGACTGAGCAGTAGAGGTAGAGATTAAGATCCGATTGAAAATT 420
QY 465 GAGCAGAGGAGCTGTAAAGTGCACAAAGCACAACAGAGATTGTAAAGGCTATTAAAG 524
|||||
Db 421 GAGCAGAGGAGCTGTAAAGTGCACAAAGCACAACAGAGATTGTAAAGGCTATTAAAG 480
QY 525 ACTTATGGGGTTCCTGTTGTAGAGGACCTTCTGAAGCAGAAAGATGTGAGCCCT 584
|||||
Db 481 ACTTATGGGGTTCCTGTTGTAGAGGACCTTCTGAAGCAGAAAGATGTGAGCCCT 540
QY 585 TTGCATTAAC 594
|||||
Db 541 TTGCATTAAC 550

RESULT 3
AT881599 586 bp mRNA linear EST 02-FEB-2000
LOCUS 606068G09.y1 606 - Ear tissue cDNA library from Schmidt lab Zea
DEFINITION
ACCESSION AT881599
VERSION AT881599.1 GI:5566733
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 586)
Walbot V.
Maize ESTs from various cDNA libraries sequenced at Stanford University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 606068 row: G column: 09.
Location/Qualifiers

FEATURES
source

1..586
/organism="Zea mays"
/cultivar="Oh1043"
/db_xref="taxon:4577"
/clone_lib="606 - Ear tissue cDNA library from Schmidt lab"
/tissue_type="mixed"
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOLR (Stratagene)"
/note="Organ: Immature ear; Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt lab"

BASE COUNT 186 a 132 c 159 g 109 t
ORIGIN

Query Match 23.9%; Score 350; DB 9; Length 586;
Best Local Similarity 99.3%; Pred. No. 1.3e-69;
Matches 550; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

20 CGCGGTTCTTGGGCCCTCCGGCTCAGCGCGCCGCCACCGCCGCGCCGAG 79
|||||
33 CGGCGTTCTTGGGCCCTCCGGCTCAGCGCGCCGCCACCGCCGCGCCGAG 92
|||||
80 ACGAGATGGGCATCAAGGTTTGACGAACGCTGGCGGACAAATGGCGCCCAAGCGCATGA 139
|||||
93 ACGAGATGGGCATCAAGGTTTGACGAACGCTGGCGGACAAATGGCGCCCAAGCGCATGA 152
|||||
140 AGGAGCAGAAAGTTCGAGAGCTACTTCGCCGCAAAATCGCGCTGACGCCGACATGAGCA 199
|||||
153 AGGAGCAGAAAGTTCGAGAGCTACTTCGCCGCAAAATCGCGCTGACGCCGACATGAGCA 212
|||||
200 TATACCAAGTTCGATGTTAGTTGGAGAGACAGCATGAAACTCTCACAAAATGAGCTG 259
|||||
213 TCTACCAAGTTCGATGTTAGTTGGAGAGACAGCATGAAACTCTCACAAAATGAGCTG 272
|||||
260 GTGAGTCACTAGTCAATTCGAGAGATGTTCAACCGGACAAATGAAATTTCTGGAAGCGG 319
|||||
273 GTGAGTCACTAGTCAATTCGAGAGATGTTCAACCGGACAAATGAAATTTCTGGAAGCGG 332
|||||
320 GAATCAAGCAGTTTATGTTTGGATGGCAAGCCTCCTGATATGAGAAACAAAGAGCTTG 379
|||||
333 GAATCAAGCAGTTTATGTTTGGATGGCAAGCCTCCTGATATGAGAAACAAAGAGCTTG 392
|||||
380 CTAAAGATCTCAAAAAGATGATGCAACCAAGATCTGACTGAGGACAGTAGAGTAG 439
|||||
393 CTAAAGATCTCAAAAAGATGATGCAACCAAGATCTGACTGAGGACAGTAGAGTAG 452
|||||
440 GAGATAAGATGATGAAATTTGAGCAAGAGACTGTAAGAGTACAAAGGACAGCA 499
|||||
453 GAGATAAGATGATGAAATTTGAGCAAGAGACTGTAAGAGTACAAAGGACAGCA 512
|||||
500 ACGAAGATGTAAGCGCTATTAAGACTTATGGGGTTCTGTTGTAAGAGCAGCTTCTG 559
|||||
513 ACGAAGATGTAAGCGCTATTAAGACTTATGGGGTTCTGTTGTAAGAGCAGCTTCTG 572
|||||
560 AAGCAGAGCAGAA 573

Db 573 AACGAGAGCAGAA 586
|||||

RESULT 4
LOCUS BE186786 467 bp mRNA linear EST 22-JUN-2000
DEFINITION 946012C08.XI 946 - tassal primordium prepared by Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION BE186786
VERSION BE186786.1 GI:8665970
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 467)
Walbot V.
Maize ESTs from various cDNA libraries sequenced at Stanford University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946012 row: C column: 08.
Location/Qualifiers

FEATURES
source

1..467
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassal primordium prepared by Schmidt lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI; Site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

BASE COUNT 108 a 110 c 96 g 153 t
ORIGIN

Query Match 22.8%; Score 334; DB 9; Length 467;
Best Local Similarity 99.7%; Pred. No. 5.6e-66;
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

886 AGATATCAATTCCTGAGAGCTGCCCTTACCAAGAGCTGAGCCTGTTCAAGAGCT 945
|||||
467 AGATATCAATTCCTGAGAGCTGCCCTTACCAAGAGCTGAGCCTGTTCAAGAGCT 408
|||||
946 AATGTCAATTCGATTCCTGAGAGCTGAAATGAGCTGACCTGATGAGAGGCTCATA 1005
|||||
407 AATGTCAATTCGATTCCTGAGAGCTGAAATGAGCTGACCTGATGAGAGGCTCATA 348
|||||
1006 AGTTCTCGTAAAGATTAATGTTTCAACGCAAGATCGGCTGACAAAGGCCATAGAGAG 1065
|||||
347 AGTTCTCGTAAAGATTAATGTTTCAATGAAATTCGGGTGACAAAGGCCATAGAGAG 288
|||||
1066 ATCAATTCCTGCAAGATTAATGCTGCAAGAGAGCTGCAAGCTCTTTTCAAGCAACT 1125
|||||
287 ATCAATTCCTGCAAGATTAATGCTGCAAGAGAGAGCTGCAAGCTCTTTTCAAGCAACT 228
|||||
1126 GCCACGACATCAGCAGCGCTAAACGAGAGAGCTTCGATTAACAGCAAGGAGGAGCT 1185
|||||
227 GCCACGACATCAGCAGCGCTAAACGAGAGAGAGCTTCGATTAACAGCAAGGAGGAGCT 168
|||||

QY 1186 GCGAACAAGAAACAAAGCTGTGGTGAAGAGAAATTAATCTTGATGCTTGATGATACAA 1245
|||||
Db 167 GCGAACAAAGAAACAAAGCTGTGGTGAAGAGAAATTAATCTTGATGATGATACAA 108
QY 1246 CTAGACTAGAAAGCGCGTGGC 1270
|||||
Db 107 CTAGACTAGAAAGCGCGTGGC 83

RESULT 5
AM562789 553 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mma sequence.
ACCESSION AM562789
VERSION AM562789.1 GI:7216667
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 553)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
Location/Qualifiers
FEATURES
source 1..553
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 171 a 111 c 129 g 142 t
ORIGIN

Query Match 21.7%; Score 317; DB 9; Length 553;
Best Local Similarity 99.7%; Pred. No. 3.2e-62;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 903 GGACTGGCTTACCACAGAGCTGACGCTTGTCAAGGAGCCTTAATGTCACATTGGATAT 962
|||||
Db 4 GGAATGGCCTTACCACAAAGCTGACGCTTGTCAAGGAGCCTTAATGTCACATTGGATAT 63
QY 963 TCTGAGCTAAATGAGCTGACCTGATGAGAGGCTTCATAGTTTCTGGTAAAGA 1022
|||||
Db 64 TCCGAGCTAAATGAGCTGACCTGATGAGAGGCTTCATAGTTTCTGGTAAAGA 123
QY 1023 TAATGGTTCAAGCAATGGGGGNCAGGCAATGAGCAATGCAATTCCTCCCAAGA 1082
|||||
Db 124 TAATGGTTCAATGAAATGGGGGNCAGGCAATGAGCAATGCAATTCCTCCCAAGA 183
QY 1083 TAAATCGTCCAGAGAGAGCTGAGTCTTTTCAAGCCAACTGCGCACATCAGACACC 1142
|||||
Db 184 TAAATCGTCCAGAGAGAGCTGAGTCTTTTCAAGCCAACTGCGCACATCAGACACC 243
QY 1143 GCTTAAAGGAGAGAGAGCTTGGATTAAGCAAGGAGGCTGGCAACAGAAACAA 1202
|||||

Db 244 GCTAAAGGAGAGAGAGCTTGGATTAAGCAAGGAGGCTGCGCAACAGAAACAA 303
QY 1203 GGCCTGGAGAAAGAAATATCTTGATGATGATGATGATGATGATGATGATGATGAT 1262
|||||
Db 304 GGCCTGGAGAAAGAAATATCTTGATGATGATGATGATGATGATGATGATGATGAT 363
QY 1263 GCGGTGGC 1270
|||||
Db 364 GCGGTGGC 371

RESULT 6
A1861468/c 470 bp mRNA linear EST 19-JUL-1999
LOCUS 614014D03.x1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mma sequence.
ACCESSION A1861468
VERSION A1861468.1 GI:5525575
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 470)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614014 row: D column: 03.
Location/Qualifiers
FEATURES
source 1..470
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOLR"
/note="Organ: root; Vector: pBluescriptII SK+; Site_1:
EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
BASE COUNT 112 a 113 c 93 g 152 t
ORIGIN

Query Match 20.2%; Score 295; DB 9; Length 470;
Best Local Similarity 99.7%; Pred. No. 2.9e-57;
Matches 345; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 965 CCTGATGAGAGGGTCTCATAGTTTCTGGTAAAGATTAATGTTCAACGAAGATCGG 1044
|||||
Db 422 CCTGATGAGAGGGTCTCATAGTTTCTGGTAAAGATTAATGTTTCAACGAAGATCGG 363
QY 1045 GTGACAAAGGCCATGAGAGAGTCAATCTGCCAAGATTAATGTCGCAAGAGAGCTC 1104
|||||
Db 362 GTGACAAAGGCCATGAGAGAGTCAATCTGCCAAGATTAATGTCGCAAGAGAGCTC 303
QY 1105 GAGTCTTTTTCACGCAAGTGCACACATCAGACCGCTAAAGGAGAGAGACTTCG 1164
|||||
Db 302 GAGTCTTTTTCACGCAAGTGCACACATCAGACCGCTAAAGGAGAGAGAGACTTCG 243
QY 1165 GATAAACAAGCAAGGAGTGGCGAACAAGAAACAAAGCTGGTGAAGAGAAATAA 1224
|||||
Db 242 GATAAACAAGCAAGGAGTGGCGAACAAGAAACAAAGGCTGGTGAAGAGAAATAA 183
QY 1225 TCTTGATGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1284
|||||

Db 182 TCTTGGATGCTTGAATGACAACTTCGACTACGAAAGACGAGGCTGCCTGATCATTGCT 123

QY 1285 TAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTTGTAAGTT 1330

Db 122 TAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTTGTAAGTT 77

RESULT 7	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AI065689/c	AI065689	456 bp mRNA linear EST 24-JUL-1998	ag91fl2.1 maize inflorescence immature ear library	AI065689	clone ag91fl2 3, mRNA sequence.	AI065689.1	GI:341096
					EST.		
					Zea mays.		
					Zea mays		

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 456)
Schutz, K., de la Bastide, M., Gnoj, L., Habermann, K., Huang, E. N.,
Parnell, L.D., Dedhia, N., Matlenssen, R. and McComble, W.R.
Expressed sequence tags from Z. mays
Unpublished (1998)
Contact: W. Richard McComble

Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel.: 516 367 8884
Tel.: 516 367 8884
Fax: 516 367 8874
Email: mccomb@cschl.org
Plate: ag91 row: f column: 12
Seq primer: M13 forward universal -21
High quality sequence stop: 456.

FEATURES	Location/Qualifiers
source	1. .456

```

/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ag91f12"
/clone_lib="maize inflorescence immature ear library"
/sex="female"
/tissue_type="immature ear"
/note="Vector: pBLUESCRIPT SK+ (X52325): Site.1: XhoI;
Site.2: EcoRI. This library is described in Schmidt, Hake,
et al., (1993) Plant Cell 5:729-737. cDNAs are
directionally cloned into the XhoI and EcoRI sites; XhoI
is near the polyA tail. Most reads from this library are
3' in direction. Additional information on this library as
well as ftp access to all sequences can be found at
http://www.cshl.org/maizegenome"

```

Query March	18.1%	Score 265;	DB 9;	Length 456;
Best Local Similarity	100.0%	Pred. No. 1.5e-50;		
Matches 265;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1048	ACAAAGGCCATGAGAAATCAATCAATCTCCAAAGATAAATCGTGCACAGGAAGACTCGAG	1107
Db	365	ACAAAGGCCATGAGAAATCAATCTCCAAAGATAAATCGTGCACAGGAAGACTTCGAG	306
Qy	1108	TCTCTTTTTCACGCCAAGCTGCCACACATCACACCGCTCAAAAACGGAGAGACTTCGGAT	1167
Db	305	TCTCTTTTTCACGCCAAGCTGCCACACATCACACCGCTCAAAAACGGAGAGACTTCGGAT	246
Qy	1168	AAAACCAAGCAAGGAGCTGCGACACAGAAACAAAGCGCTGTGGAAAGAGAAATAACT	1227
Db	245	AAAACCAAGCAAGGAGCTGCGACACAGAAACAAAGCGCTGTGGAAAGAGAAATAACT	186
Qy	1228	TGGATGCTTGATGTACACACTACGACTACGAAAGCAGCGGTGGCGTATCACTTCGCTTAG	1287

Db	185	TGGATGCTTGATGTACACATTCGACACTACGAAAGCAGCGGTGGCGCTGATCACTTGGCTTAG	126
Qy	1288	ATTATTAACTCCCTGTTTAACTC	1312
Db	125	ATTATTAACTCCCTGTTTAACTC	101

RESULT	8
BG837708	
LOCUS	BG837708
DEFINITION	BG837708 901 bp mRNA linear EST 25-MAY-2001 Zm10_01f08_A Zm10_AAFC_ECCRC_Fusarium_graminearum_corn_silk Zea mays CDNA clone Zm10_01f08, mRNA sequence.
ACCESSION	BG837708
VERSION	BG837708.1 GI:14204031
KEYWORDS	EST,
SOURCE	Zea mays.
ORGANISM	Zea mays

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 901)	Harris, L.J., Balcerzak, M., Allard, S., Sparano, A., Courroux, P., De Moors, A., Hattori, J. I., Ouellet, T., Robert, L.S., Singh, J.A., Sprott, D. and Tinker, N.A.	Expressed Sequence Tags from Maize Silk Six Hours After Silk Channel Inoculation with <i>Fusarium graminearum</i>

JOURNAL COMMENT

Unpublished (2001)
Contact: Harris, Linda J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6
CANADA
Tel: (613) 759-1314
Fax: (613) 759-6566
Email: lharris1@em.agr.ca.

FEATURES	Location/Qualifiers
source	1. .901

```

/organism="Zea mays"
/cultivar="C0388"
/db_xref="taxon:4577"
/clone="Zm10_01i08"
/clone_id="Zm10_AAFc_ECORC_Fusarium_graminearum_corn_silk"
"
/tissue_type="Silk"
/dex_stage="4-5 days post-silk emergence"
/note="Vector: Bluescript SK+/XhoI-ECOR1; Site_1: ECOR1; Site_2: XhoI; Field-grown corn was silk channel-inoculated in the morning (~10 am) with 1 ml of a macroconidial suspension (500,000 spores/ml) of Fusarium graminearum and silk channels were collected and immediately frozen in liquid nitrogen 6 hours later. RNA was extracted from silk tissue between 1 cm below and above the inoculation point in the silk channel, RNA from five silk channels was pooled."

```

Query Match	Score	DB 10:	Length
18.0%	263	10	901

Matches	603;	Conservative	0;	Mismatches	2;	Indels	2;	Gaps	2;
QY	676	CCAGTTC	CAAGAAATACCTGTGATGGAATTTGATGTGGCCAGGTTTGGAGAGCCTT	735					
Db	72	CCAGTTCC	AGAAATACCTGTGATGGAATTTGATGTGGCCAGGTTTGGAGAGCCTT	131					
QY	736	GAATCACC	ATGAGACCACTTCATTGATTTGTGCATCCTGTGTGATGTGACTAATTTGAT	795					
Db	132	GAACAC	CACATGAGACCACTTCATTGATTTGTGCATCCTGTGTGATGTGACTAATTTGAT	191					
QY	796	AGCATCA	AAAGTATTCGGGGGGCAACACACTCTGAAACTTATTTCGTCACATGGGTCCATA	855					
Db	192	AGCATCA	AAAGTATTCGGGGGGCAACACACTCTGAAACTTATTTCGTCACATGGGTCCATA	251					

QY 856 GAAGACATCTTGAGAAATCTTAATAAGACAGATATCAATTC-TGAGACATGGCCCTTA 914
|||||
Db 252 GAAAGCATCTTGAGAAATCTTAATAAGACAGATATCAATTCCTGTGAGACATGGCCCTTA 311
QY 915 CC-AAGAGCTCGACGCTTGTTCAGGAGCCTTAATGTTCATATGATATTCCTGAGCTAA 973
|||||
Db 312 CCAAGAGATCGACGCTTGTTCAGGAGCCTTAATGTTCATATGATATTCCTGAGCTAA 371
QY 974 AATGACTGCACTGATGAGAGAGGCTCTCAATATTCCTGTGTAAGATATGTTTCA 1033
|||||
Db 372 AATGACTGCACTGATGAGAGAGGCTCTCAATATTCCTGTGTAAGATATGTTTCA 431
QY 1034 ACAGAACTCGGATGACAAAGGCCATAGAGATCAATTCGCCAAGATATTCGTCG 1093
|||||
Db 432 ATGAGAGTGGGATGACAAAGGCCATAGAGATCAATTCGCCAAGATATTCGTCG 491
QY 1094 AAGGAAGCTCGAGTCTCTTTTTCAGCCCACTGCGCACCATCATGACGACCGCTAAAGCGA 1153
|||||
Db 492 AAGGAAGCTCGAGTCTCTTTTTCAGCCCACTGCGCACCATCATGACGACCGCTAAAGCGA 551
QY 1154 AGGAGACTGGATAAAGCAAGGCAAGCTGGGAAACAAGAAAGAGCTGTGAA 1213
|||||
Db 552 AGGAGACTGGATAAAGCAAGGCAAGCTGGGAAACAAGAAAGAGCTGTGAA 611
QY 1214 AGAAGAAATATCTTGATGCTGTATGATACACTACGACTACGAAAGCAGCGTGGCGTG 1273
|||||
Db 612 AGAAGAAATATCTTGATGCTGTATGATACACTACGACTACGAAAGCAGCGTGGCGTG 671
QY 1274 ATCACTT 1280
|||||
Db 672 ATCACTT 678

RESULT 9
AW559173/c 474 bp mRNA linear EST 07-MAR-2000
LOCUS 660065H06.xl 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW559173
VERSION AW559173.1 GI:7204640
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS 1 (bases 1 to 474)
TITLE Walbot,V.
JOURNAL Maize ESTs from various cDNA libraries sequenced at Stanford
COMMENT Unpublished (1999)
CONTACT: Walbot V
DEPARTMENT OF Biological Sciences
STANFORD UNIVERSITY
855 California Ave, Palo Alto, CA 94304, USA
TEL: 650 723 2227
FAX: 650 725 8221
EMAIL: walbot@stanford.edu
PLATE: 660065 row: H column: 06.
LOCATION/Qualifiers
1. .474
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcORI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amle Franklin."

BASE COUNT 113 a 114 c 92 g 154 t 1 others
ORIGIN
Query Match 17.0%; Score 249; DB 9; Length 474;
Best Local Similarity 99.4%; Pred. No. 5; Gc-47;
Matches 349; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 920 AAGCTGACGGTTGTCAGAGGCTTAATGTCACATTCGATATTCCTGAGCTAAATGGA 979
|||||
Db 474 AAGCTGACGGTTGTTTCAGAGGCTTAATGTCACATTCGATATTCCTGAGCTAAATGGA 415
QY 980 CTGCACTGATGAGAGGCTCTCATAGTTTCTGTGTAAGATATGTTTCAACGAG 1039
|||||
Db 414 CTGCACTGATGAGAGGCTCTCATAGTTTCTGTGTAAGATATGTTTCAATGAG 355
QY 1040 ATCGGTGACAAAGGCCATAGAGAAATCAATCTGCCAATAATATTCGCGAAGGA 1099
|||||
Db 354 ATCGGTGACAAAGGCCATAGAGAAATCAATCTGCCAATAATATTCGCGAAGGA 295
QY 1100 GACTCGAGTCTCTTTTTCAGGCACTGCGCACCATCATGACGACCGCTAAAGCAGAGA 1159
|||||
Db 294 GACTCGAGTCTCTTTTTCAGGCACTGCGCACCATCATGACGACCGCTAAAGCAGAGA 235
QY 1160 CTTCGATAAACAAGCAAGGCACTGCGAACAAGAAACAAAGGCTGTGGAAGAAGA 1219
|||||
Db 234 CTTCGATAAACAAGCAAGGCACTGCGAACAAGAAACAAAGGCTGTGGAAGAAGA 175
QY 1220 AATATCTTGATGCTGTATGATACACTACGACTACGAAACACGCGTGGC 1270
|||||
Db 174 AATATCTTGATGCTGTATGATACACTACGACTACGAAACACGCGTGGC 124

RESULT 10
AW000375 532 bp mRNA linear EST 08-SEP-1999
LOCUS 614014D03.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW000375
VERSION AW000375.1 GI:5847296
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS 1 (bases 1 to 532)
TITLE Walbot,V.
JOURNAL Maize ESTs from various cDNA libraries sequenced at Stanford
COMMENT Unpublished (1999)
CONTACT: Walbot V
DEPARTMENT OF Biological Sciences
STANFORD UNIVERSITY
855 California Ave, Palo Alto, CA 94304, USA
TEL: 650 723 2227
FAX: 650 725 8221
EMAIL: walbot@stanford.edu
PLATE: 614014 row: D column: 03.
LOCATION/Qualifiers
1. .532
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOLR"
/note="Organ: root; Vector: pBluescriptII SK+; Site_1:
EcORI; Site_2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"

BASE COUNT 173 a 103 c 129 g 127 t
ORIGIN

```

Query Match          16.7%; Score 245; DB 9; Length 532;
Best Local Similarity 100.0%; Pred. No. 4.1e-46;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 953 CATTGATATTCCTCGTCAAAATGAGTCGCACCTGATGAGAGGGCTTCATAGTTTC 1012
|
|
|
Db 288 CATTGATATTCCTCGTCAAAATGAGTCGCACCTGATGAGAGGGCTTCATAGTTTC 347
|
|
|
QY 1013 TGGTAAAGATATGATGTTTCAACGAGATCGGGTGACAAAGCCATAGAGAAATCAAT 1072
|
|
|
Db 348 TGGTAAAGATATGATGTTTCAACGAGATCGGGTGACAAAGCCATAGAGAAATCAAT 407
|
|
|
QY 1073 CTGCAGAGATTAATGCTGCGAAGAGACTGAGTCTTTTCAAGCCAATGCCACCA 1132
|
|
|
Db 408 CTGCAGAGATTAATGCTGCGAAGAGACTGAGTCTTTTCAAGCCAATGCCACCA 467
|
|
|
QY 1133 CATCAGCACCCTAAAGGAGAGACTTGGATTAACCAAGCAGCGCTGCCAACA 1192
|
|
|
Db 468 CATCAGCACCCTAAAGGAGAGACTTGGATTAACCAAGCAGCGCTGCCAACA 527
|
|
|
QY 1193 AGAAA 1197
|
|
|
Db 528 AGAAA 532
|
|
|

RESULT 11
AM562517 475 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.X2.660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AM562517
VERSION AM562517.1 GI:7216395
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE 1 (bases 1 to 475)
COMMENT Walbot, V.
JOURNAL Maize ESTs from various cDNA libraries sequenced at Stanford
COMMENT University
Unpublished (1999)
CONTACT Walbot V
DEPARTMENT Department of Biological Sciences
STANFORD UNIVERSITY
855 California Ave, Palo Alto, CA 94304, USA
TEL: 650 723 2227
FAX: 650 725 8221
EMAIL: walbot@stanford.edu
LOCATION/Qualifiers
FEATURES
source
1. 475
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 124 a 110 c 94 g 147 t
ORIGIN
Query Match          16.5%; Score 241; DB 9; Length 475;
Best Local Similarity 99.7%; Pred. No. 3.5e-45;
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 979 ACTGCACCTGATGAGAGGGCTCTCATAGTTTCTGTAAAGATATGTTTCAACGAA 1038
|
|
|

```

```

Db 475 ACTGCACCTGATGAGAGGGCTCTCATAGTTTCTGTAAAGATATGTTTCAATGAA 416
|
|
|
QY 1039 GATCGGGTGACAAAGCCCATAGAGATCAATTCGCCAAGATTAATGTCGCAAGCA 1098
|
|
|
Db 415 GATCGGGTGACAAAGCCCATAGAGATCAATTCGCCAAGATTAATGTCGCAAGCA 356
|
|
|
QY 1099 AGACTGAGTCTTTTTCAGCCCACTGCCACCAATCAGCACCGCTAAACGGAAGAG 1158
|
|
|
Db 355 AGACTGAGTCTTTTTCAGCCCACTGCCACCAATCAGCACCGCTAAACGGAAGAG 296
|
|
|
QY 1159 ACTTCGATTAACCAAGCAGCGCTGCCACCAAGAAAAGAGTGTGTGAAAGAG 1218
|
|
|
Db 295 ACTTCGATTAACCAAGCAGCGCTGCCACCAAGAAAAGAGTGTGTGAAAGAG 236
|
|
|
QY 1219 AATTAATCTGATGCTTGTATGACACTGCACTACGAAAGCAGCGGTGC 1270
|
|
|
Db 235 AATTAATCTGATGCTTGTATGACACTGCACTACGAAAGCAGCGGTGC 184
|
|
|

RESULT 12
A1065546 376 bp mRNA linear EST 24-JUL-1998
LOCUS ag88e02.x1 maize inflorescence immature ear library Zea mays cDNA
DEFINITION clone ag88e02 3', mRNA sequence.
ACCESSION A1065546
VERSION A1065546.1 GI:3340953
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE 1 (bases 1 to 376)
COMMENT Parnell, L.D., Dedhia, N., Martienssen, R. and McCombie, W.R.
JOURNAL Expressed sequence tags from Z. mays
UNPUBLISHED (1998)
CONTACT: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
COLD SPRING HARBOR LABORATORY
PO BOX 100, COLD SPRING HARBOR, NY 11724, USA
TEL: 516 367 8884
FAX: 516 367 8874
EMAIL: mcombie@cshl.org
PLATE: ag88 row: e column: 02
Seq primer: M13 forward universal -21
High quality sequence stop: 376.
LOCATION/Qualifiers
FEATURES
source
1. 376
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="ag88e02"
/clone_lib="maize inflorescence immature ear library"
/sex="female"
/tissue_type="immature ear"
/note="Vector: PHUESCRIPT SK+ (X52325); Site_1: XhoI;
Site_2: EcoRI; This library is described in Schmidt, Hake,
et al., (1993) Plant Cell 5:729-737. cDNAs are
directionally cloned into the XhoI and EcoRI sites; XhoI
is near the polyA tail. Most reads from this library are
3' in direction. Additional information on this library as
well as ftp access to all sequences can be found at
http://www.cshl.org/maizegenome"
BASE COUNT 79 a 99 c 65 g 131 t 2 others
ORIGIN
Query Match          15.4%; Score 225; DB 9; Length 376;
Best Local Similarity 99.6%; Pred. No. 1.5e-41;
Matches 275; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1016 TAAAGATTAATGTTTCAACGAGATCGGGTGACAAAGCCATAGAGAGATCAATCTG 1075
|
|
|

```

Db 376 TAAAGATTAATGTTTCACGAGATCGGTTGACAAAGGCCATAGAGAGATCAATCTG 317

QY 1076 CCAAGATTAATGTCGCAAGAGACCTCGCTCTTTTCAAGCCAACTGCCACCAT 1135
|||||

Db 316 CCAAGATTAATGTCGCAAGAGACCTCGCTCTTTTCAAGCCAACTGCCACCAT 257
|||||

QY 1136 CAGCACCCTAAACGAGAGACTTCGATTAACAGACGACCTCGCAACAAGA 1195
|||||

Db 256 CAGCACCCTAAACGAGAGACTTCGATTAACAGACGACCTCGCAACAAGA 197
|||||

QY 1196 AACCAAGGCTGTGGAAGAGAAATATCTTGATGATCACTACGACTAC 1255
|||||

Db 196 AACCAAGGCTGTGGAAGAGAAATATCTTGATGATCACTACGACTAC 137
|||||

QY 1256 GAAGCAGCGGTGGGATGATCCTGCTTAGATTA 1291
|||||

Db 136 GAAGCAGCGGTGGGATGATCCTGCTTAGATTA 101
|||||

RESULT 13
AM288831/c 470 bp mRNA linear EST 16-JAN-2000
LOCUS 707010F11.y2 707 - Mixed adult tissues from Walbot lab (Sk) Zea
DEFINITION
ACCESSION AM288831
VERSION mays cDNA, mRNA sequence.
KEYWORDS
SOURCE EST.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 470)
REFERENCE
AUTHORS Walbot, V
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: F column: 11.
FEATURES
source
Location/Qualifiers
1..470
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_idb="707 - Mixed adult tissues from Walbot lab (Sk)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 102 a 110 c 108 g 150 t

ORIGIN

Query Match 15.3%; Score 224; DB 9; Length 470;
Best Local Similarity 99.6%; Pred. No. 2.2e-41;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 962 TTCCTGAGCTAAATGAGACTGACCTGATGAGAGGCTCTCATAGTTTCTGTAAG 1021
|||||

Db 429 TTCCTGAGCTAAATGAGACTGACCTGATGAGAGGCTCTCATAGTTTCTGTAAG 370
|||||

QY 1022 ATAAATGTTTCACGAGATCGGTTGACAAAGCCATAGAGAGATCAATCTGCCAGA 1081
|||||

Db 369 ATAAATGTTTCACGAGATCGGTTGACAAAGGCCATAGAGAGATCAATCTGCCAGA 310

QY 1082 ATAAATGTTTCACGAGAGAGATCGCTCTTTTCAAGCCAACTGCCACCAT 1141
|||||

Db 309 ATAAATGTTTCACGAGAGAGATCGCTCTTTTCAAGCCAACTGCCACCAT 250
|||||

QY 1142 CGCTAAACGAGAGACTTCGATTAACAGACGACCTCGCAACAACA 1201
|||||

Db 249 CGCTAAACGAGAGACTTCGATTAACAGACGACCTCGCAACAACA 190
|||||

QY 1202 AGGCTGTGGAAGAGAAATATCTTGATGCTT 1236
|||||

Db 189 AGGCTGTGGAAGAGAAATATCTTGATGCTT 155
|||||

RESULT 14
AM288784 414 bp mRNA linear EST 16-JAN-2000
LOCUS 707010C02.x5 707 - Mixed adult tissues from Walbot lab (Sk) Zea
DEFINITION
ACCESSION AM288784
VERSION mays cDNA, mRNA sequence.
KEYWORDS
SOURCE EST.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 414)
REFERENCE
AUTHORS Walbot, V
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: C column: 02.
FEATURES
source
Location/Qualifiers
1..414
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_idb="707 - Mixed adult tissues from Walbot lab (Sk)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site:1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 120 a 82 c 101 g 109 t 2 others

ORIGIN

Query Match 13.1%; Score 192; DB 9; Length 414;
Best Local Similarity 99.0%; Pred. No. 3.4e-34;
Matches 392; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 649 CCAGGTTCTCTGCTCATTTAATGATGCCAAGTCCAAAGAAATATCTGATGGAATT 708
|||||

Db 6 CCAGGTTCTCTGCTCATTTAATGATGCCAAGTCCAAAGAAATATCTGATGGAATT 65
|||||

QY 709 GATTTGCCAAGGTTTGGAGAGCTTGAATCCATGACCAAGTTCATTTGATTTGTC 768
|||||

Db 66 GATTTGCCAAGGTTTGGAGAGCTTGAATCCATGACCAAGTTCATTTGATTTGTC 125
|||||

QY 769 ATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 828
|||||

Db 126 ATCTGCTGATGATGCTATTGTGATGATCAAAAGCTATCGGGGGGCAACAGCTCTG 185
QY 829 AACTTATTCGTACATGGGCTCCATAGAAAGCATCTTGGAGAATCTTAATTAAGACAGA 888
Db 186 AACTTATTCGTACATGGGCTCCATAGAAAGCATCTTGGAGAATCTTAATTAAGACAGA 245
QY 889 TATCAATTCCTGAGGAGCTGGCCCTTACCAAGAAAGCTGACGCTTGTTCAGAGACCTAAT 948
Db 246 TACCAATTCCTGAGGAGCTGGCCCTTACCAAGAAAGCTGACGCTTGTTCAGAGACCTAAT 305
QY 949 GTCAATTCGATATTCCTGAGCTAAATAGACATGCATGAGGAGGCTTCATAGT 1008
Db 306 GTACATTCGATATTCCTGAGCTAAATAGACATGCATGAGGAGGCTTCATAGT 365
QY 1009 TTCCTGTTAAAGATATGCTTTCACAGCAAGATCG 1044
Db 366 TTCCTGTTAAAGATATGCTTTCACAGCAAGATCG 401

RESULT 15
AM562788/c 225 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.X5 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.

ACCESSION AM562788
VERSION AM562788.1 GI:7216666
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.

FEATURES
source Location/Qualifiers
1..225
/organism="Zea mays"
/cultivar="Oh1043"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="X10LR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."

BASE COUNT 40 a 53 c 45 g 87 t
ORIGIN

Query Match 12.8%; Score 187; DB 9; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.5e-33;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 GAAGATGGGTGACAAAGGCGATAGAGAATCAATCTGCCAAGATAATCTCGCAA 1095
Db 187 GAAGATGGGTGACAAAGGCGCGATAGAGAATCAATCTGCCAAGATAATCTCGCAA 128
QY 1096 GGAAGACTCGAGTCCTTTTCAAGCCAACTGCGCCACATCAGACCGCTAAAGCGGAG 1155
Db 127 GGAAGACTCGAGTCCTTTTCAAGCCAACTGCGCCACATCAGACCGCTAAAGCGGAG 68

QY 1156 GAGACTTCGGATTAACACAGCAGCTGCGAACAAGAAACAAAGCTGTGGAAG 1215
Db 67 GAGACTTCGGATTAACACAGCAGCTGCGAACAAGAAACAAAGCTGTGGAAG 8
QY 1216 AAGAAAT 1222
Db 7 AAGAAAT 1

RESULT 16
AM288831 470 bp mRNA linear EST 16-JAN-2000
LOCUS 707010F11.Y2 707 - Mixed adult tissues from Walbot lab (SK) Zea
DEFINITION mays cDNA, mRNA sequence.

ACCESSION AM288831
VERSION AM288831.1 GI:6695753
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
UNIVERSITY

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: F column: 11.

FEATURES
source Location/Qualifiers
1..470
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
)"

/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 102 a 110 c 108 g 150 t
ORIGIN

Query Match 10.1%; Score 148; DB 9; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.2e-24;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 717 CAAGTTTGGAGAGGCTGTGACATCACCATGACAGCTCATTTATTTGTCATCTGTG 776
Db 14 CAAGTTTGGAGAGGCTGTGACATCACCATGACAGCTCATTTATTTGTCATCTGTG 73

QY 777 TGGATGTGACTATTTGATAGCATCAAGATATGCGGGGCAACACCTGAACTTAT 836
Db 74 TGGATGTGACTATTTGATAGCATCAAGATATGCGGGGCAACACCTGAACTTAT 133

QY 837 TCGTCAACATGGGTCCATAGCAAGCATC 864
Db 134 TCGTCAACATGGGTCCATAGCAAGCATC 161

RESULT 17
AM562518/c 363 bp mRNA linear EST 10-MAR-2000
LOCUS AM562518

```

DEFINITION 660065H06.x3 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.
ACCESSION AWS62518
VERSION AWS62518.1 GI:7216396
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 363)
REFERENCE 1
AUTHORS Walbot,V.
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
FEATURES
source
1. .363
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XL0LR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: ECORI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Ame Franklin."
BASE COUNT 96 a 82 c 74 g 111 t
ORIGIN
Query Match 8.9%; Score 130; DB 9; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.7e-20;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1141 CCGCTAAAGGAGAGACTTCGATTAACAACAAGCAGCTCGGACAGAAAACA 1200
|||||
Db 301 CCGCTAAAGGAGAGACTTCGATTAACAACAAGCAGCTCGGACAGAAAACA 242
|||||
QY 1201 AAGCTGTGGAGAGAAATTAATCTTGATGCTGATGACACTACGACGAAG 1260
|||||
Db 241 AAGCTGTGGAGAGAAATTAATCTTGATGCTGATGACACTACGACGAAG 182
|||||
QY 1261 CAGCGGTGGC 1270
|||||
Db 181 CAGCGGTGGC 172
|||||

RESULT 18
BF727781 232 bp mRNA linear EST 09-JAN-2001
LOCUS 1000052F12.x2 1000 - Unigene 1 from Maize Genome Project Zea mays
DEFINITION cDNA, mRNA sequence.
ACCESSION BF727781
VERSION BF727781.1 GI:12045642
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 232)
REFERENCE 1
AUTHORS Walbot,V.
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.
FEATURES
source
1. .363
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XL0LR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: ECORI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Ame Franklin."
BASE COUNT 96 a 82 c 74 g 111 t
ORIGIN

```

```

JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1000052 row: F column: 12.
FEATURES
source
1. .232
/organism="Zea mays"
/db_xref="dbEST:707010C02.y1"
/clone_lib="1000 - Unigene 1 from Maize Genome Project"
/note="This library represents the unique ESTs found in
the first round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
libraries 486, 487, 496, 603, 605, 606, 614, 618, 660, 683
, 687, 707, and 945. Contigs were assembled using TIGR's
CAP program and a representative EST from each contig was
selected for the Unigene set. All singlets were also
selected."
BASE COUNT 64 a 42 c 57 g 69 t
ORIGIN
Query Match 7.0%; Score 103; DB 10; Length 232;
Best Local Similarity 100.0%; Pred. No. 3.9e-14;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 755 TCATTGATTTGTGCATCTGTGTGATGTGACTATTGTGATGATCAAGGTATCGGG 814
|||||
Db 55 TCATTGATTTGTGCATCTGTGTGATGTGACTATTGTGATGATCAAGGTATCGGG 114
|||||
QY 815 GGCAACAGCTCTGAACCTATTGCTCACATGGGTCACATGA 857
|||||
Db 115 GGCAACAGCTCTGAACCTATTGCTCACATGGGTCACATGA 157
|||||

RESULT 19
A1834484 553 bp mRNA linear EST 02-FEB-2000
LOCUS 606068G09.x1 606 - Ear tissue cDNA library from Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION A1834484
VERSION A1834484.1 GI:5468693
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 553)
REFERENCE 1
AUTHORS Walbot,V.
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 606068 row: G column: 09.
FEATURES
source
1. .553
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="606 - Ear tissue cDNA library from Schmidt
lab"
/tissue_type="mixed"

```

```
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOLR (Stratagene)"
/notes="Organ: Immature ear; Vector: pBR-CMW; Site_1: EcoRI
; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
lab"

BASE COUNT      140 a      112 c      106 g      195 t

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 68; DB 9; Length 553;
Pred. No. 3,9e-12;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 GAAGATCGGTCGCAAGGCCATGAGAGATCAATCTCCAGAGTAATATGTCGCAA 1095
|||||
Db 512 GAAGATCGGTCGCAAGGCCATGAGAGATCAATCTCCAGAGTAATATGTCGCAA 453
|||||

QY 1096 GGAAGACTGAGTCCTTTTCAAGCCACTGCC 1128
|||||
Db 452 GGAAGACTGAGTCCTTTTCAAGCCACTGCC 420
|||||

RESULT 20
LOCUS      AM147048      126 bp      mRNA      linear      EST 03-NOV-1999
DEFINITION 707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
ACCESSION  AM147048
VERSION     AM147048.1  GI:6194944
KEYWORDS   EST.
SOURCE      Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 126)
REFERENCE   1
  AUTHORS   Walbot V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 707010 row: C column: 02.

FEATURES
  source
    1..126
    /organism="Zea mays"
    /cultivar="W23"
    /db_xref="taxon:4577"
    /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
    )"
    /tissue_type="tassel, kernel, silk, husk, root, leaf"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Organ: tassel, kernel, silk, husk, root, leaf;
    Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
    differentiated maize tissues from an active Mutator
    plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
    husk, root, leaf). Unidirectionally cloned."

BASE COUNT      29 a      31 c      26 g      40 t

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 68; DB 9; Length 126;
Pred. No. 3,7e-06;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 797 GCATCAAGGTATCGGGGCAAAAGCTGTGAACCTATTGTCACATGTCGCTAG 856
|||||
Db 1 GCATCAAGGTATCGGGGCAAAAGCTGTGAACCTATTGTCACATGTCGCTAG 60
|||||
```

```
QY 857 AAGCATC 864
|||||
Db 61 AAGCATC 68
|||||

RESULT 21
LOCUS      AM147048/c      126 bp      mRNA      linear      EST 03-NOV-1999
DEFINITION 707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
ACCESSION  AM147048
VERSION     AM147048.1  GI:6194944
KEYWORDS   EST.
SOURCE      Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 126)
REFERENCE   1
  AUTHORS   Walbot V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 707010 row: C column: 02.

FEATURES
  source
    1..126
    /organism="Zea mays"
    /cultivar="W23"
    /db_xref="taxon:4577"
    /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
    )"
    /tissue_type="tassel, kernel, silk, husk, root, leaf"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Organ: tassel, kernel, silk, husk, root, leaf;
    Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
    differentiated maize tissues from an active Mutator
    plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
    husk, root, leaf). Unidirectionally cloned."

BASE COUNT      29 a      31 c      26 g      40 t

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 65; DB 9; Length 126;
Pred. No. 1,8e-05;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1172 CAAGCAAGCAGCTGCGACACAGAAACAAAGCTGTGCAAGAGAAATATCTTGA 1231
|||||
Db 126 CAAGCAAGCAGCTGCGACACAGAAACAAAGCTGTGCAAGAGAAATATCTTGA 67
|||||

QY 1232 TGCTT 1236
|||||
Db 66 TGCTT 62
|||||

RESULT 22
LOCUS      A1947478      225 bp      mRNA      linear      EST 19-AUG-1999
DEFINITION 614047B01.x1 614 - root cDNA library from Walbot lab Zea mays cDNA,
ACCESSION  A1947478
VERSION     A1947478.1  GI:5739683
KEYWORDS   EST.
SOURCE      Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

REFERENCE 1 (bases 1 to 225)
Spermatophyta; Magnoliophyta, Liliopsida; Poales; Poaceae; PACC
clade; Panicoidae; Andropogoneae; Zea.

AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614047 row: B column: 01.

FEATURES
Source
1..225
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLDLR"
/note="Organ: root; Vector: pBluescriptII SK+; Site_1:
ECORI; Site_2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM) "
Location/Qualifiers

BASE COUNT 66 a 54 c 40 g 65 t
ORIGIN

Query Match 3.9%; Score 57; DB 9; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1185 TGGCAGACAGAAACAAAGCTGCTGCAAGAAATATCTTGATGATCT 1241
|||||
Db 225 TGGCAGACAGAAACAAAGCTGCTGCAAGAAATATCTTGATGATCT 169

RESULT 23
LOCUS BG406288 346 bp mRNA linear EST 28-NOV-2001
DEFINITION sac29u02.y1 Gm-cl051 glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl051-3819 5' similar to TR:Q9SWS9 Q9SWS9 RIBOSOMAL PROTEIN
S26.; mRNA sequence.
ACCESSION BG406288
VERSION BG406288.1 GI:13312637
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 346)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritters, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Putative full length read
vector to vector length is 347 This clone is available through:
ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL

TITLE
JOURNAL
COMMENT

FEATURES
Source
1..346
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-3819"
/clone_lib="Gm-cl051"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knap of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

BASE COUNT 107 a 49 c 73 g 117 t
ORIGIN

Query Match 2.8%; Score 41; DB 10; Length 346;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1423 TTTCAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1463
|||||
Db 304 TTTCAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 344

RESULT 24
LOCUS AL121294/c 468 bp mRNA linear EST 25-FEB-2000
DEFINITION DKFZp762L016.t1 762 (synonym: hme12) Homo sapiens cDNA clone
DKFZp762L016 5', mRNA sequence.
ACCESSION AL121294
VERSION AL121294.1 GI:5927295
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 468)
Bloecher, H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Bloecher, et al.)
Unpublished (1999)
Contact: Bloecher H

TITLE
JOURNAL
COMMENT

FEATURES
Source
1..468
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp762L016"
/clone_lib="762 (synonym: hme12)"
/tissue_type="melanoma (Mewo cell line)"
/dev_stage="adult"
/lab_host="DH10B"


```

BASE COUNT      100 a      110 c      104 g      154 t
ORIGIN
Query Match      2.8%; Score 41; DB 9; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1423 TTGAAGTAAAAA...1463
|||||
DB 68 TTGAAGTAAAAA...28

RESULT 25
LOCUS      B1761155      642 bp      mRNA      linear      EST 25-SEP-2001
DEFINITION 603043646F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5184131 5',
            mRNA sequence.
ACCESSION   B1761155
VERSION     B1761155.1 GI:157523733
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 642)
REFERENCE   NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM11459 row: k column: 12
            High quality sequence stop: 383.
FEATURES
            source
            1..642
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5184131"
            /clone_lib="NIH_MGC_116"
            /lab_host="DH10B"
            /note="Organ: pooled colon, kidney, stomach; Vector:
            pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 3 colons, age 26 yo male, 49 yo
            female, 71 yo male colon; 46 yo male kidney, and pool of 2
            stomachs, 62 yo male and 70 yo female. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.4 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            023. Note: this is a NIH MGC Library."
BASE COUNT      234 a      148 c      151 g      109 t
ORIGIN
Query Match      2.8%; Score 41; DB 10; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1423 TTGAAGTAAAAA...1463
|||||
DB 269 TTGAAGTAAAAA...309

RESULT 26
LOCUS      B1760815      822 bp      mRNA      linear      EST 25-SEP-2001

```

```

DEFINITION 603043646F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5184154 5',
            mRNA sequence.
ACCESSION   B1760815
VERSION     B1760815.1 GI:15752393
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 822)
REFERENCE   NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM11459 row: 1 column: 11
            High quality sequence start: 3
            High quality sequence stop: 370.
FEATURES
            source
            1..822
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5184154"
            /clone_lib="NIH_MGC_116"
            /lab_host="DH10B"
            /note="Organ: pooled colon, kidney, stomach; Vector:
            pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 3 colons, age 26 yo male, 49 yo
            female, 71 yo male colon; 46 yo male kidney, and pool of 2
            stomachs, 62 yo male and 70 yo female. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.4 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            023. Note: this is a NIH MGC Library."
BASE COUNT      301 a      205 c      211 g      105 t
ORIGIN
Query Match      2.8%; Score 41; DB 10; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1423 TTGAAGTAAAAA...1463
|||||
DB 268 TTGAAGTAAAAA...308

RESULT 27
LOCUS      AM307211      202 bp      mRNA      linear      EST 02-DEC-2001
DEFINITION sf54b11.y1 Gm-c1009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1009-3766 5', mRNA sequence.
ACCESSION   AM307211
VERSION     AM307211.1 GI:6719564
KEYWORDS
SOURCE      soybean.
ORGANISM    Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
            1 (bases 1 to 202)
REFERENCE   Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corveill,V., Khanna
            A., Bolla,B., Maitra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

```

TITLE
JOURNAL
COMMENT
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com
Seq primer: -40RP from Gibco.

FEATURES

Source

1. 202
Location/Qualifiers
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl009-3766"
/clone_lib="Gm-cl009"
/lab_host="XLI0-Gold"

/note="Vector: pBluescript II XR, Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from entire roots of 2-month-old 'Williams' plants that were greenhouse grown in 5-gallon pots. To suppress nodulation, Black Gold All-Purpose potting soil was supplemented with: 0.36g/L available phosphoric acid (P205), 20mg/L urea N, 0.16g/L S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.5mg/L Mn, 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following nutrients in a slow-release form (Osmocote): 0.165g/L ammonia N, 0.185g/L nitrate N, 0.35g/L available phosphoric acid, and 0.35g/L soluble potash. No nodules were visible on the roots at harvest. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated. Stratagene's first-strand synthesis primer was used [GAGAGAGAGAGAGAGAGACTACTGCGAG(T)-18]. After second-strand synthesis, the cDNA ends were 'polished' with clone Pfu DNA polymerase, ligated to EcoRI adaptors, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 400bp cutoff, using a SizeSeph 400 Spin column from Pharmacia. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) that had been digested with EcoRI and XhoI, and phosphorylated. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts. This library was constructed by Dr. Paul Keim and Dr. Virginia Correll."

BASE COUNT 102 a 22 c 22 g 56 t
ORIGIN

Query Match 2.7%; Score 40; DB 9; Length 202;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 TTGAAGTAAAAA 1463
Db 127 TTGAAGTAAAAA 166

RESULT 28
AI373782/c 203 bp mRNA linear EST 16-FEB-1999
LOCUS AI373782
DEFINITION q254906.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2030746 3',
mRNA sequence.
ACCESSION AI373782
VERSION AI373782.1 GI:4153648

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 203)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/RLNL at:
www.bio.lnlnl.gov/bbrp/image/image.html
Insert length: 304 Std Error: 0.00
Seq primer: -40RP from Gibco.

FEATURES

source

1. 203
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2030746"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"

/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 59 a 30 c 28 g 86 t
ORIGIN

Query Match 2.7%; Score 40; DB 9; Length 203;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1424 TTGAAGTAAAAA 1463
Db 42 TTGAAGTAAAAA 3

RESULT 29
AA678858/c

LOCUS AA678858 294 bp mRNA linear EST 02-DEC-1997
DEFINITION ah05403.s1 Gessler Wilms tumor Homo sapiens cDNA clone
IMAGE:1155724 3', mRNA sequence.
AA678858
AA678858.1 GI:2659380
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 294)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maita, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

TITLE
JOURNAL
COMMENT

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m13 fwd. ET from Amer sham
 High quality sequence stop: 216.
 Location/Qualifiers

FEATURES

source

1. 294

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1155724"

/clone_lib="Gessler Wilms tumor"

/sex="pooled (6)"

/lab_host="DH10B"

/note="vector: pSPORT1; Site_1: SalI; Site_2: NotI; RNA

was prepared from a pool of 6 anonymous Wilms' tumor RNAs.

RNA was prepared by acid-phenol, followed by one round of

oligo dt selection. cDNA library preparation was with

the BRL/Life Tech. Superscript Plasmid system. An

oligo-dt NotI primer for first strand synthesis generated

ggcgccgccctn at the 3' end of the clones. A 5' SalI

adaptor was used with sequence 5'-gtcagccacgcgcgcg-3'.

Resulting cDNAs were size selected (average size 2 kb),

NotI digested, and ligated into NotI/SalI-cut pSPORT1.

Library was constructed by Dr. Manfred Gessler."

BASE COUNT

89 a 46 c 49 g 110 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 40; DB 9; Length 294;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 TTGAAGTAAAAA
 DB 47 TTGAAGTAAAAA
 1463

DEFINITION

LOCUS BG301426 409 bp mRNA linear EST 09-MAY-2001

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

The library was constructed by Brandi Chiappelli and Dr. James

McCarter (jchapel@estewatson.wustl.edu & jmcarter@estewatson.wustl.edu) at

The vector to vector length is 410
 Seq primer: SL1 primer.
 Location/Qualifiers

FEATURES

source

1. 409

/organism="Strongyloides ratti"

/db_xref="taxon:34506"

/clone_lib="Strongyloides ratti L1 SL1 Topol Chiappelli

McCarter"

/dev_stage="L1"

/lab_host="DH10B"

/note="vector: PCR11-TOPO; SL1-oligo(dt) PCR-based

library. Strongyloides_ratti_L1 cDNA PCR products of size

>400 nucleotides containing SL1 on the 5' end and oligo(dt

) on the 3' end were non-directionally cloned into

PCR11-TOPO (Invitrogen) following the TOPO TA cloning

protocol. The cDNA insert can be excised by digestion with

ECOR1. The library was constructed by Brandi Chiappelli and

Dr. James McCarter at Washington University, St. Louis.

Nematodes were provided by Dr. Mark Viney of the

University of Bristol, UK."

BASE COUNT

154 a 69 c 73 g 113 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 40; DB 10; Length 409;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 TTGAAGTAAAAA
 DB 370 TTGAAGTAAAAA
 1463

LOCUS BM378668/c 411 bp mRNA linear EST 16-JAN-2002

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2001)

Contact: Patrick S. Schnable

Schnable Laboratory

Iowa State University

G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA

Tel: 515-294-0975

Fax: 515-294-2299

Email: schnable@iastate.edu

Individual basecall and confidence value were assigned using the

Phred software(<http://www.phrap.org/>). Overall sequence quality

assessment and vector trimming was conducted using the Lucy

software (version 1.16s, <http://www.ligr.org/softlab/>). Lucy

parameters were set to ensure an overall trimmed quality of 97.5%

or better without any vector fragments in the chosen high quality

region of each sequence. Low-quality bases between the poly-T and
 the high-quality region were replaced with N's to serve as spacers
 using a Perl program (est_process.pl), written by Dr. Hui-Hsien
 Chou.

PCR PRIMERS

FORWARD: T3 (ATT AAC CCT CAC TAA AC)

BACKWARD: T7-1 (AA TAC GAC TCA CTA TAG)

Seq primer: UNLV (GTA AAA CGA CCG CCA GT)

PolyA=yes.

Location/Qualifiers

1. 411

source

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1424 TCGAAGTAAAAA 1463
 Db 192 TTGAAGTAAAAA 231

RESULT 34
 AM568974
 LOCUS 109 bp mRNA linear EST 03-DEC-2001
 DEFINITION s173909.y1 Gm-cl031 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl031-377 5', mRNA sequence.
 ACCESSION AM568974
 VERSION AM568974.1 GI:7233569
 KEYWORDS EST
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 109)
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna,
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,
 R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: cu@resgen.com
 High quality sequence stop: 67.
 Location/Qualifiers
 1..109
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl031-377"
 /clone_lib="Gm-cl031"
 /tissue_type="Williams seedlings, minus the cotyledons"
 /lab_host="D10B"
 /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from whole 'Williams' seedlings, minus the cotyledons,
 which were propagated on paper towels with distilled water
 for 5 days, incubated at 40 degrees C for 1 hour. The
 cotyledons were removed and the remaining tissue was
 flash-frozen in liquid nitrogen. StrataGene's cDNA
 Synthesis Kit (catalog number 200401) was used to
 synthesize the cDNA. First-strand synthesis was performed
 with 5-methyl dCTP, hence the ligated cDNA was
 hemimethylated. A modification of StrataGene's
 first-strand synthesis primer was used. An 'anchor'
 nucleotide (V=A, C, or G) was added to the 3' end of the
 primer [GAGACAGACAGACAGACACTGCTCCAG(T)18V] to anchor
 the primer at the 5' end of the poly(A) tract. After
 second strand synthesis, the cDNA ends were filled in with
 cloned Pfu DNA, ligated to EcoRI adapters and
 subsequently phosphorylated. The cDNA was then
 precipitated and redissolved in sterile, RNase-
 dnase-free water. The XhoI site within the first-strand
 synthesis primer was then restricted by digestion with
 XhoI from Promega (400/ul); all XhoI sites in the cDNA
 would be protected by their hemimethylated status. The
 cDNA constructs were size-fractionated with a 500 bp

BASE COUNT 67 a 7 c 14 g 21 t

ORIGIN

Query Match 2.7%; Score 39; DB 9; Length 109;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA 1463
 Db 52 TGAAGTAAAAA 90

RESULT 35
 BI319195
 LOCUS 150 bp mRNA linear EST 23-JUL-2001
 DEFINITION 949026H11.y1 949 - Juvenile leaf and shoot cDNA from Steve Moose
 zea mays cDNA, mRNA sequence.
 ACCESSION BI319195
 VERSION BI319195.1 GI:14997152
 KEYWORDS EST.
 SOURCE zea mays.
 ORGANISM zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 150)
 Walbot, V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 8227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 949026 row: H column: 11.
 Location/Qualifiers
 1..150
 /organism="Zea mays"
 /cultivar="M64A"
 /db_xref="taxon:4577"
 /clone_lib="949 - Juvenile leaf and shoot cDNA from Steve
 Moose"
 /tissue_type="Immature leaf primordium and vegetative
 meristem"
 /dev_stage="4 stages from 3-13 days after imbibing"
 /lab_host="E. coli XL10R"
 /note="Organ: Juvenile vegetative shoots; Vector:
 pAD-GAL4-2.1; Site_1: EcoRI; Site_2: XhoI; Equal amounts
 of total RNA by weight from 4 tissue sources (see below)
 were pooled, polyA+ RNA isolated, and cDNA synthesized for
 EcoRI (5') and XhoI (3') directional cloning into lambda
 HybriZap vector from StrataGene. Tissue Sources: 1. Whole
 shoots 3 days after sowing/imbibing in wet soil. 2. Basal
 1.5 cm shoots 6 days after sowing - includes yellow
 portions of developing leaves 1-5, primordia from 6-8, and
 the vegetative apex. 3. Non-green portions of developing
 leaves 4-5 and the vegetative apex, including adult leaf
 primordia. 9 days after sowing. 4. Partially expanded and
 greening leaves 4-5 at 13 days after sowing."

BASE COUNT 86 a 17 c 14 g 33 t

ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 150;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 95 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 133

RESULT 36
BF749502 187 bp mRNA linear EST 10-JAN-2001
LOCUS IL0-BN0422-121000-500-b06 BN0422 Homo sapiens CDNA, mRNA sequence.
DEFINITION BF749502
ACCESSION BF749502.1 GI:12076178
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 187)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/ICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-BN0422-
121000-500-b06&t3=2000-10-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 186.
Location/Qualifiers
1. 187
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0422"
/dev_stage="Adult"
/note="Organ: breast, normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 78 a 25 c 36 g 48 t
ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 187;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 142 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 180

RESULT 37

BF749503 201 bp mRNA linear EST 10-JAN-2001
LOCUS IL0-BN0422-121000-500-b09 BN0422 Homo sapiens CDNA, mRNA sequence.
DEFINITION BF749503
ACCESSION BF749503
VERSION BF749503.1 GI:12076179
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 201)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/ICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL0&t2=IL0-BN0422-
121000-500-b09&t3=2000-10-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 34
High quality sequence stop: 200.
Location/Qualifiers
1. 201
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0422"
/dev_stage="Adult"
/note="Organ: breast, normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

JOURNAL MEDLINE
COMMENT

FEATURES

source

BASE COUNT 82 a 27 c 41 g 51 t
ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 201;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
|||||
Db 158 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 196

RESULT 38
BF764516 226 bp mRNA linear EST 12-JAN-2001
LOCUS R01-CS0068-011100-011-g03 CS0068 Homo sapiens CDNA, mRNA sequence.
DEFINITION BF764516
ACCESSION BF764516.1 GI:12112416
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 226)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and Simpson, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPERP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC1&f2=RC1-CS0068-011100-011-g03&t3=2000-11-01&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 25
 High quality sequence stop: 186.

FEATURES
 source location/Qualifiers
 1..226
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CS0068"
 /dev_stage="Adult"
 /note="Organ: colon_est; Vector: puc18; site:1: Sma1; site:2: Sma1: A mini-library was made by cloning products derived from ORESTES PCR (O.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 122 a 27 c 33 g 44 t

ORIGIN

Query Match 2.7%; Score 39; DB 10; Length 226;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA
 Db 132 TGAAGTAAAAA

RESULT 39
 AM234203 320 bp mRNA linear EST 02-DEC-2001
 LOCUS sf22d11.y1 Gm-c1028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION Gm-c1028-718 5', mRNA sequence.
 ACCESSION AM234203
 VERSION AM234203.1 GI:6566560
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 1 (bases 1 to 320)
 Shoemaker, R., Kelm, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, R., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Rifter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 Public Soybean EST Project

JOURNAL Unpublished (1999)
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cut@resgen.com
 Insert Length: 394 Std Error: 0.00
 High quality sequence stop: 306.

FEATURES
 source location/Qualifiers
 1..320
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-718"
 /clone_lib="Gm-c1028"
 /issue_type="roots of 'Supernod' plants"
 /lab_host="DH10B"
 /note="Vector: pBluescript II XR; Site:1: EcoRI; site:2: XhoI; The mRNA was isolated from roots of Glycine max 'Supernod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. StrataGene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A-C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGACGACGACGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 110 a 38 c 66 g 106 t

ORIGIN

Query Match 2.7%; Score 39; DB 9; Length 320;
 Best Local Similarity 100.0%; Pred. No. 6.5;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1425 TGAAGTAAAAA
 Db 275 TGAAGTAAAAA

RESULT 40
 A2953545 365 bp DNA linear GSS 27-APR-2001
 LOCUS A2953545/c
 DEFINITION 2M0218A19R Mouse 10kb plasmid UUGCM library Mus musculus genomic
 ACCESSION A2953545
 VERSION A2953545.1 GI:13824772
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
AUTHORS	1 (bases 1 to 365)		
TITLE	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duvall,B., Haml,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,R., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.		
JOURNAL COMMENT	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000) Contract: Robert B. Weiss University of Utah Genome Center Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0218 row: A column: 19 Seq primer: CACACAGGAACACGCTAAGACC Class: plasmid ends High quality sequence stop: 365.		
FEATURES	Location/Qualifiers 1..365 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC2M0218A19" /clone_id="Mouse 10kb plasmid UUGC2M library" /sex="female" /lab_host="E. coli strain XL10-Gold, T1-resistant, F-" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMDA2 (g14732114[gbl AF129072.1]), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."		
BASE COUNT	85 a 65 c 40 g 175 t		
ORIGIN			
Query Match	2.7%; Score 39; DB 12; Length 365;		
Best Local Similarity	100.0%; Pred. No. 6;		
Matches 39; Conservative 0; Mismatches 0;	Indels 0; Gaps 0;		
CY 1425	TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463		
Db 150	TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 112		
RESULT 41			
AA121171/c			
LOCUS	AA121171 431 bp mRNA linear EST 19-MAY-1997		
DEFINITION	2188905..51 Stratagene colon (#937204) Homo sapiens cDNA clone IMAGE:511736 3' similar to gp:M65066 CAMP-DEPENDENT PROTEIN KINASE TYPE I-ALPHA REGULATORY (HUMAN);, mRNA sequence.		
ACCESSION	AA121171		
VERSION	AA121171.1 GI:1678714		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		

REFERENCE	Makaryova; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 431)
AUTHORS	Hallier,L., Lennon,G., Becker,M., Bernaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins, R., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore, B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,D., Trevasakis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)
TITLE	97044478
JOURNAL	
MEDLINE	
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: estelw@stunl.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 1257 Std Error: 0.00 Seq primer: -40M13 fwd. from Amersham High quality sequence stop: 93. Location/Qualifiers 1.431 /organism="Homo sapiens" /db_xref="GDB:3844231" /db_xref="taxon:9606" /clone="IMAGE:511736" /clone_lib="Stratagene colon (#937204)" /tissue_type="tumor" /cell_line="T84 carcinoma cell line" /lab_host="SOLR cells (kanamycin resistant)" /note="Organ: colon; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: oligo gp. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' CTCGCGTTTTTTTTTTTTTTT 3' "
BASE COUNT	84 a 103 c 125 g 107 t 12 others
ORIGIN	
Query Match	2.7%; Score 39; DB 9; Length 431;
Best Local Similarity	100.0%; Pred. No. 5.4;
Matches	39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1425 TGAAGTAAAAA.....AAAAAAAAAAAAA 1463
Db	46 TGAAGTAAAAA.....AAAAAAAAAAAAA 8
RESULT 42	
AA121150/c	AA121150 432 bp mRNA linear EST 19-MAY-1997
LOCUS	
DEFINITION	Z186d07.s1 Stratagene colon (#937204) Homo sapiens cDNA clone IMAGE:511693 3' similar to gp:DU0761 PROTEASOME COMPONENT C5 (HUMAN); mRNA sequence.
ACCESSION	AA121150
VERSION	AA121150
KEYWORDS	AA121150.1 GI:1678693
SOURCE	EST.
ORGANISM	human.
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 432)	
REFERENCE	Hallier,L., Lennon,G., Becker,M., Bernaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins, R., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore, B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,D., Trevasakis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)
TITLE	
JOURNAL	
MEDLINE	
COMMENT	

MEDLINE 97044478
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1006 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 354.

FEATURES

source

```
1. 432
/organism="Homo sapiens"
/db_xref="GDB:3844188"
/db_xref="taxon:9606"
/clone_lib="IMAGE:511693"
/clone_lib="Stratagene colon (#937204)"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/notes="Organ: colon; Vector: pBluescript SK-; Site: 1:
EcoRI; Site: 2: XhoI; Cloned unidirectionally. Primer:
Oligo df. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor
sequence: 5' GAATCGGCGACGAG-3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3'."
```

BASE COUNT 105 a 111 c 79 g 137 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 5.4;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1425 TGAAGTAAAAAATTTTTTTTTTTTTTTTTTTT 1463

Db 39 TGAAGTAAAAAATTTTTTTTTTTTTTTTTTTT 1

RESULT 43

BM357517/c

444 bp mRNA linear EST 30-JAN-2002

LOCUS

BM357517 Triphysaria versicolor root-tp, early DMBQ-induced

DEFINITION

transcript cDNA library Triphysaria versicolor cDNA, mRNA sequence.

ACCESSION

BM357517

VERSION

BM357517.1 GI:18426469

KEYWORDS

EST.

SOURCE

Triphysaria versicolor.

ORGANISM

Triphysaria versicolor.

REFERENCE

1 (bases 1 to 444)

AUTHORS

Matvienko, M., Torres, M.J. and Yoder, J.I.

TITLES

Transcriptional responses to the hemiparasitic plant Triphysaria

JOURNAL

Plant Physiol. 127 (1), 272-282 (2001)

MEDLINE

21437952

COMMENT

Contact: Yoder, J.I.

FEATURES

1. 444

source

/organism="Triphysaria versicolor"

/db_xref="taxon:64093"

/clone_lib="Triphysaria versicolor root-tp, early

DMBQ-induced transcript cDNA library"

/tissue_type="root-tips ~5mm length"

/dev_stage="3-4 weeks growth"

/lab_host="E. coli"

/note="Vector: pCR2.1 TA Cloning System, Invitrogen,

Carlsbad, CA; PCR-based suppression subtractive

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

hybridization cDNA library"

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 ||||||||||||||||||||||||||||||||||||||||
 DB 42 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4

RESULT 45
 AL501532/c 593 bp mRNA linear EST 04-JAN-2001
 LOCUS AL501532 Hordeum vulgare Barke roots Hordeum vulgare cDNA clone
 DEFINITION HM041040 3', mRNA sequence.
 ACCESSION AL501532
 VERSION AL501532.1 GI:12027747
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
 ; Triticeae; Hordeum.
 1 (bases 1 to 593)
 Michael, W., Weschke, W., Pleissner, K.-P. and Graner, A.
 EST sequencing and analysis in barley
 Unpublished (2000)
 Contact: Michael, W.
 Institute for Plant Genetics and Crop Plant Research
 Corrensstr.3, D-06466 Gatersleben, Germany
 Email: michael@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de
 Seq primer: Mjuni primer for 3'end.
 Location/Qualifiers
 1..593
 /organism="Hordeum vulgare"
 /cultivar="Barke"
 /db_xref="taxon:4513"
 /clone_1lb="HM041040"
 /clone_1lb="Hordeum vulgare Barke roots"
 /tissue_type="roots"
 /lab_host="XIOLR"
 /note="Vector: Plasmid pRK-CMV; Site.1: EcoRI; Site.2:
 XhoI; mRNA was made from roots of spring barley variety
 'Barke', a high quality malting variety. Roots were grown
 for two days on filter paper at room temperature. Cloning
 sites: EcoRI (5'-end of cDNA) and XhoI (3'-end of cDNA).
 NOTE: Due to a cloning artefact caused by the kit, in most
 cases the EcoRI site is NOT present, as well as the EcoRI
 adapter. Average insert size is 1 kb. Sequence trimming:
 Vector sequences and sequence ends were trimmed from the
 5'- and 3'-end until a 50 bp window contains less than two
 ambiguities. The maximum length was set to 700 bp"
 1 others

BASE COUNT 165 a 122 c 114 g 191 t
 ORIGIN

Query Match 2.7%; Score 39; DB 9; Length 593;
 Best Local Similarity 100.0%; Pred.No. 4.5;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1425 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1463
 ||||||||||||||||||||||||||||||||||||||||
 DB 42 TGAAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4

Search completed: November 5, 2002, 16:21.48
 Job time : 1287.78 secs

```

XX RESULT 1
XX ID AAY95307
XX AAY95307 standard; Protein; 379 AA.
XX AC AAY95307;
XX DT 12-SEP-2000 (first entry)
XX DE Maize Rad2/FEN-1 protein.
XX MA Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
XX KW endonuclease; exonuclease; DNA repair; gene targeting.
XX OS Zea mays.
XX WO200036109-A1.
XX PN
XX PD 22-JUN-2000.
XX PE 16-NOV-1999; 99WO-US27147.
XX PR 15-DEC-1998; 98US-0112332.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Mahajan PB;
XX WI; 2000-452026/39.
XX DR N-PSDB; AAA27923.
XX MAIZE RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
XX recombination and repair in transgenic plants, e.g. for gene targeting
XX and the production of male sterile plants -
XX

```

PS Claim 11, Page 71-72; 85pp; English.
XX
CC The present sequence is that of maize RAD2/FEN-1, as deduced from
CC a cDNA clone (see AAA27923) isolated from maize line B73 immature
CC ear tissue. Rad2/FEN-1 is a structure specific endonuclease which
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC can be expressed in transgenic plant cells using conventional
CC methods. The protein is involved in the regulation of DNA repair
CC and recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 379 AA;

Query Match 100.0%; Score 379; DB 21; Length 379;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIGLTKRLADNAPKAMKEQKFESEYFGRKIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
DB 1 MGIGLTKRLADNAPKAMKEQKFESEYFGRKIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
QY 61 VTSHLQGFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
DB 61 VTSHLQGFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
QY 121 KDAIEKLSKRTVYVTRQHNEDCKRLRLMGVPVVEAPSEAECAALCINDKVFVAASED 180
DB 121 KDAIEKLSKRTVYVTRQHNEDCKRLRLMGVPVVEAPSEAECAALCINDKVFVAASED 180
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLELELTMDQFIDLCILCGDYCDSIK 240
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLELELTMDQFIDLCILCGDYCDSIK 240
QY 241 GIGGOTALKLIRQHGSIESTILENLNKDRQIPEDMPYQEARRLFKEPNVTLDIPELKMTA 300
DB 241 GIGGOTALKLIRQHGSIESTILENLNKDRQIPEDMPYQEARRLFKEPNVTLDIPELKMTA 300
QY 301 PDEGLISFLVKNDFNEDRVTAKIEIKSAKKNSSQGRLESEFFKPTATTSAPLKRKETS 360
DB 301 PDEGLISFLVKNDFNEDRVTAKIEIKSAKKNSSQGRLESEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAANKKTKAGGKK 379
DB 361 DKTSKAANKKTKAGGKK 379

RESULT 2
AA95309
ID AA95309 standard; Protein; 379 AA.
XX
AC AA95309;
XX
DT 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 protein.
XX
KM Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
XX endonuclease; exonuclease; DNA repair; gene targeting.
OS Zea mays.
XX
XX WO200036109-A1.
PN
XX
PD 22-JUN-2000.

XX
PF 16-NOV-1999; 99WO-US27147.
XX
PR 15-DEC-1998; 98US-0112332.
XX
PA (PION-) PIONEER HI-BRED INT. INC.
XX
PI Mahajan PB;
XX
DR WPI: 2000-452026/39.
XX
N-PSDB: AAA27925.
XX
CC The present sequence is that of maize RAD2/FEN-1, as deduced from
CC a cDNA clone (see AAA27925) derived from maize line W23 tassel
CC polyA RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC can be expressed in transgenic plant cells using conventional
CC methods. The protein is involved in the regulation of DNA repair
CC and recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 379 AA;

Query Match 100.0%; Score 379; DB 21; Length 379;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIGLTKRLADNAPKAMKEQKFESEYFGRKIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
DB 1 MGIGLTKRLADNAPKAMKEQKFESEYFGRKIAVDASMSIYQFLIVGRTGMETLTNEAGE 60
QY 61 VTSHLQGFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
DB 61 VTSHLQGFNRTIRLLEAGIKPVYVFDGKPPDMKKOELAKRYSKRDATKDLTEAVEVGD 120
QY 121 KDAIEKLSKRTVYVTRQHNEDCKRLRLMGVPVVEAPSEAECAALCINDKVFVAASED 180
DB 121 KDAIEKLSKRTVYVTRQHNEDCKRLRLMGVPVVEAPSEAECAALCINDKVFVAASED 180
QY 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLELELTMDQFIDLCILCGDYCDSIK 240
DB 181 MDSLTFGAPRFLRLHMDPSSKKIPVMEFDYAKVLELELTMDQFIDLCILCGDYCDSIK 240
QY 241 GIGGOTALKLIRQHGSIESTILENLNKDRQIPEDMPYQEARRLFKEPNVTLDIPELKMTA 300
DB 241 GIGGOTALKLIRQHGSIESTILENLNKDRQIPEDMPYQEARRLFKEPNVTLDIPELKMTA 300
QY 301 PDEGLISFLVKNDFNEDRVTAKIEIKSAKKNSSQGRLESEFFKPTATTSAPLKRKETS 360
DB 301 PDEGLISFLVKNDFNEDRVTAKIEIKSAKKNSSQGRLESEFFKPTATTSAPLKRKETS 360
QY 361 DKTSKAANKKTKAGGKK 379
DB 361 DKTSKAANKKTKAGGKK 379

RESULT 3
AA95308

```

ID  AAY95308 standard; Protein; 379 AA.
XX
XX  AAY95308;
AC
XX  12-SEP-2000 (first entry)
DT
XX  Maize Rad2/FEN-1 protein.
DE
XX  Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
KW  endonuclease; exonuclease; DNA repair; gene targeting.
XX
XX  Zea mays.
OS
XX  WO200036109-A1.
PN
XX  22-JUN-2000.
PD
XX  16-NOV-1999; 99WO-US27147.
PF
XX  15-DEC-1998; 98US-0112332.
PR
XX  (PION-) PIONEER HI-BRED INT INC.
PA
XX  Mahajan PB;
PI
XX  WPI: 2000-452026/39.
DR  N-PSDB: AAA27924.
XX
XX  Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT  recombination and repair in transgenic plants, e.g. for gene targeting
PT  and the production of male sterile plants -
XX
XX  Example 1; Page 74; 85pp; English.
XX
XX  The present sequence is that of maize Rad2/FEN-1, as deduced from
CC  a cDNA clone (see AAA27924) derived from maize line B73 seedling
CC  tissue RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC  under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC  can be expressed in transgenic plant cells using conventional
CC  methods. The protein is involved in the regulation of DNA repair
CC  and recombination in plant systems and therefore may be used for
CC  improving gene targeting during further recombinant DNA protocols
CC  involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC  in DNA replication and nucleotide excision and repair reactions.
CC  The exolytic activity is involved in double strand break repair and
CC  end joining. The protein is also useful in strand exchange
CC  reactions during homologous recombination. These functions may be
CC  useful in gene targeting and in the production of male sterile
CC  plants. The efficacy of gene targeting can be improved by the
CC  overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC  be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX  Sequence 379 AA:
SO
Query Match 84.7%; Score 321; DB 21; Length 379;
Best Local Similarity 100.0%; Pred. No. 9.5e-303;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGKIGLTKLADNAPKAMKEQKFESYFGKRIADASMSIYQFLIVGRTGMEITLNAGE 60
DB 1 MGKIGLTKLADNAPKAMKEQKFESYFGKRIADASMSIYQFLIVGRTGMEITLNAGE 60
QY 61 VTSHLQGMFNRTIRLEAGJIKPYVVPDGRPPDMKKQELAKRSKRDATKDLFEAVVGD 120
DB 61 VTSHLQGMFNRTIRLEAGJIKPYVVPDGRPPDMKKQELAKRSKRDATKDLFEAVVGD 120
QY 121 KDAIEKLSKRTVAVTROHNECKRLLRLMGVPVVEADSEAEACALCINDKYFAVASED 180
DB 121 KDAIEKLSKRTVAVTROHNECKRLLRLMGVPVVEADSEAEACALCINDKYFAVASED 180
QY 181 MDSITFEAPRFLRLMLPSSSKKIPVMEFDVAKYLEELTMDQFTIDICICGCGYCSIK 240
DB 181 MDSITFEAPRFLRLMLPSSSKKIPVMEFDVAKYLEELTMDQFTIDICICGCGYCSIK 240

```

```

QY 241 GIGGOTALKIRHSGSIESILENLNKRQYQIPEDMPYQEARRLKEPNVTLIDIPELKMTA 300
DB 241 GIGGOTALKIRHSGSIESILENLNKRQYQIPEDMPYQEARRLKEPNVTLIDIPELKMTA 300
QY 301 PDEGLISFLVKDNGFNEDRV 321
DB 301 PDEGLISFLVKDNGFNEDRV 321
RESULT 4
AAY95310
ID  AAY95310 standard; Protein; 379 AA.
XX
XX  AAY95310;
AC
XX  12-SEP-2000 (first entry)
DT
XX  Maize Rad2/FEN-1 protein.
DE
XX  Maize: Rad2/FEN-1; transgenic plant; male sterile plant;
KW  endonuclease; exonuclease; DNA repair; gene targeting.
XX
XX  Zea mays.
OS
XX  WO200036109-A1.
PN
XX  22-JUN-2000.
PD
XX  16-NOV-1999; 99WO-US27147.
PF
XX  15-DEC-1998; 98US-0112332.
PR
XX  (PION-) PIONEER HI-BRED INT INC.
PA
XX  Mahajan PB;
PI
XX  WPI: 2000-452026/39.
DR  N-PSDB: AAA27926.
XX
XX  Maize Rad2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT  recombination and repair in transgenic plants, e.g. for gene targeting
PT  and the production of male sterile plants -
XX
XX  Example 1; Page 79-80; 85pp; English.
XX
XX  The present sequence is that of maize Rad2/FEN-1, as deduced from
CC  a cDNA clone (see AAA27926) derived from maize line B73 endosperm
CC  RNA. Rad2/FEN-1 is a structure specific endonuclease which
CC  under certain conditions also acts as an exonuclease. Rad2/FEN-1
CC  can be expressed in transgenic plant cells using conventional
CC  methods. The protein is involved in the regulation of DNA repair
CC  and recombination in plant systems and therefore may be used for
CC  improving gene targeting during further recombinant DNA protocols
CC  involving plants. Rad2/FEN-1 endonucleolytic activity is essential
CC  in DNA replication and nucleotide excision and repair reactions.
CC  The exolytic activity is involved in double strand break repair and
CC  end joining. The protein is also useful in strand exchange
CC  reactions during homologous recombination. These functions may be
CC  useful in gene targeting and in the production of male sterile
CC  plants. The efficacy of gene targeting can be improved by the
CC  overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC  be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX  Sequence 379 AA:
SO
Query Match 73.4%; Score 278; DB 21; Length 379;
Best Local Similarity 99.7%; Pred. No. 4.8e-261;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGKIGLTKLADNAPKAMKEQKFESYFGKRIADASMSIYQFLIVGRTGMEITLNAGE 60
DB 1 MGKIGLTKLADNAPKAMKEQKFESYFGKRIADASMSIYQFLIVGRTGMEITLNAGE 60

```

[illegible]

XX	RESULT 5
XX	AAB96644
XX	ID AAB96644 standard; Protein; 343 AA.
XX	AC
XX	AAB96644;
DT	29-OCT-2001 (first entry)
XX	DE Putative P. abyssi nuclease.
XX	Hyperthermophilic archaeon; hyperthermophilic protein.
XX	KW Pyrococcus abyssi.
OS	Pyrococcus abyssi.
PN	FR2792651-A1.
PD	27-OCT-2000.
PF	21-APR-1999; 99FR-0005034.
PR	21-APR-1999; 99FR-0005034.
PA	(CNRS) CNRS CENT NAT RECH SCI.
PI	(IFRE-) IFREMER INST FR RECH EXPL MER.
PI	Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
DR	Querellou J, Weissenbach J, Saurin W, Hellig R;
DR	WPI: 2001-126236/14.
PT	New nucleotide sequences isolated from Pyrococcus abyssi encode
PT	proteins useful in industry -
PS	Claim 7: Pages 1394-1395; 1657pp; French.
CC	The present invention relates to the genomic sequence of Pyrococcus
CC	abyssi (see AA886431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC	a hyperthermophilic archæon, which is isolated from deep-sea
CC	hydrothermal vents. The present invention is one such P. abyssi protein.
CC	The proteins of the present invention have various potential industrial
CC	uses, since the proteins are stable at very high temperatures, some up to
CC	110 degrees centigrade.
CC	Note: This patent is in the same patent family as WO200065062, which
CC	contains additional sequences as shown in AAB99132-AA99143,
CC	AAH75903-AAH75920 and AAG6436.
XQ	Sequence 343 AA:

```
Query Match      4.0%   Score 15; DB 22; Length 343;
Best Local Similarity 100.0%; Pred.No. 5.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0;
QY 77 EAQIKPVIYEDGKRP 91
    |||
Db 70 EAQIKPVIYEDGKRP 84
```

RESULT 6	
ABB63960	
ID	ABB63960 standard; Protein; 385 AA.
XX	
AC	ABB63960;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 18672.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide pharmaceutical.
XX	
OS	Drosophila melanogaster.

XX	MO2001/1042-A2.
XX	
PD	27-SEP-2001.
XX	
PE	23-MAR-2001; 2001WO-US09231.
XX	
FR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
XX	N-PSTDB; ABL08063.
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Disclosure; SEQ ID NO 18672; 21pp + Sequence Listing; English.
CC	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.

[illegible]

```

XX AC AAW92505;
XX DT 23-APR-1999 (first entry)
XX DE Mouse FEN-1 protein.
XX KM FEN-1; mouse; flap endonuclease; detection; diagnosis; carcinogen;
XX KN neoplasma; antineoplastic agent; cleavage.
XX OS Mus sp.
XX PN US5874283-A.
XX PD 23-FEB-1999.
XX PF 30-MAY-1995; 95US-0455968.
XX PR 30-MAY-1995; 95US-0455968.
XX PS (HARR/) HARRINGTON J J.
XX PA (HSIEH/) HSIEH C.
XX PA (LIEBER/) LIEBER M R.
XX PI Harrington JJ, Hsieh C, Lieber MR;
XX DR WPI: 1999-179985/15.
XX DR N-PSDB; AAX02108.
XX PT DNA encoding flap endonuclease polypeptides - useful for producing
XX PT e.g. recombinant polypeptides
XX PS Claim 1: Fig 2A; 58pp; English.
XX CC This sequence represents a mouse FEN-1 (flap endonuclease) protein. This
XX CC protein can be used in methods for detecting a pathological condition in
XX CC a patient, for diagnostic purposes, for screening for antineoplastic
XX CC agents and carcinogens, for diagnostic staging of neoplasia, for
XX CC producing recombinant flap endonuclease for use as research or
XX CC diagnostic reagents, for producing antibodies reactive with the novel
XX CC polypeptides, for producing transgenic nonhuman animals expressing the
XX CC novel polypeptides encoded by a transgene. The invention also provides
XX CC novel molecular cloning techniques and reagents involving cleavage of
XX CC a flap or nick with a flap endonuclease.
XX SQ Sequence 377 AA:
XX
XX Query Match 3.4%; Score 13; DB 20; Length 377;
XX Best Local Similarity 100.0%; Pred. No. 0.00052;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 GIKPVYVFDGKPP 91
DB 76 GIKPVYVFDGKPP 88

```

```

RESULT 8
AAW92508
ID AAW92508 standard; Protein: 378 AA.
XX
XX AAW92508;
XX
XX 23-APR-1999 (first entry)
XX DE Human FEN-1 protein.
XX
XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
XX KM neoplasma; antineoplastic agent; cleavage.
XX OS Homo sapiens.
XX PN US5874283-A.
XX DR

```

```

PD 23-FEB-1999.
XX 30-MAY-1995; 95US-0455968.
XX PF 30-MAY-1995; 95US-0455968.
XX PR 30-MAY-1995; 95US-0455968.
XX PS (HARR/) HARRINGTON J J.
XX PA (HSIEH/) HSIEH C.
XX PA (LIEBER/) LIEBER M R.
XX PI Harrington JJ, Hsieh C, Lieber MR;
XX DR WPI: 1999-179985/15.
XX DR N-PSDB; AAX02111.
XX PT DNA encoding flap endonuclease polypeptides - useful for producing
XX PT e.g. recombinant polypeptides
XX PS Disclosure; Fig 5A-B; 58pp; English.
XX CC This sequence represents a human FEN-1 (flap endonuclease) protein. This
XX CC protein can be used in methods for detecting a pathological condition in
XX CC a patient, for diagnostic purposes, for screening for antineoplastic
XX CC agents and carcinogens, for diagnostic staging of neoplasia, for
XX CC producing recombinant flap endonuclease for use as research or
XX CC diagnostic reagents, for producing antibodies reactive with the novel
XX CC polypeptides, for producing transgenic nonhuman animals expressing the
XX CC novel polypeptides encoded by a transgene. The invention also provides
XX CC novel molecular cloning techniques and reagents involving cleavage of
XX CC a flap or nick with a flap endonuclease.
XX SQ Sequence 378 AA:
XX
XX Query Match 3.4%; Score 13; DB 20; Length 378;
XX Best Local Similarity 100.0%; Pred. No. 0.00052;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 GIKPVYVFDGKPP 91
DB 76 GIKPVYVFDGKPP 88

```

```

RESULT 9
AAW92504
ID AAW92504 standard; Protein: 380 AA.
XX
XX AAW92504;
XX
XX 23-APR-1999 (first entry)
XX DE Human FEN-1 protein.
XX
XX FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
XX KM neoplasma; antineoplastic agent; cleavage.
XX OS Homo sapiens.
XX PN US5874283-A.
XX PD 23-FEB-1999.
XX PF 30-MAY-1995; 95US-0455968.
XX PR 30-MAY-1995; 95US-0455968.
XX PS (HARR/) HARRINGTON J J.
XX PA (HSIEH/) HSIEH C.
XX PA (LIEBER/) LIEBER M R.
XX PI Harrington JJ, Hsieh C, Lieber MR;
XX DR WPI: 1999-179985/15.
XX DR N-PSDB; AAX02107.

```

XX DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 PS Claim 1; Fig 1A; 58pp; English.
 CC This sequence represents a human FEN-1 (flap endonuclease) protein. This
 CC protein can be used in methods for detecting a pathological condition in
 CC a patient, for diagnostic purposes, for screening for antineoplastic
 CC agents and carcinogens, for diagnostic staging of neoplasia, for
 CC producing recombinant flap endonuclease for use as research or
 CC diagnostic reagents, for producing transgenic nonhuman animals expressing the
 CC polypeptides, for producing transgenic nonhuman animals expressing the
 CC novel polypeptides encoded by a transgene. The invention also provides
 CC novel molecular cloning techniques and reagents involving cleavage of
 CC a flap or nick with a flap endonuclease.
 CC
 SQ Sequence 380 AA;
 Query Match 3.4%; Score 13; DB 20; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.00052;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 79 GIKPVYFDGKPP 91
 |||||||||
 Db 78 GIKPVYFDGKPP 90
 RESULT 10
 AAM59950
 ID AAM59950 standard; Protein; 325 AA.
 AC AAM59950;
 XX
 DT 22-DEC-1998 (first entry)
 XX
 DE Amino acid sequence of the structure specific nuclease 2.
 XX
 KW Tag mutant gene; thermostable; nuclease; mutant; DNA polymerase;
 KM bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
 XX
 OS Synthetic.
 XX
 PN WO9823774-A1.
 XX
 PD 04-JUN-1998.
 XX
 PF 26-NOV-1997; 97WO-US21783.
 XX
 PR 02-DEC-1996; 96US-0758314.
 PR 29-NOV-1996; 96US-0757653.
 XX
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Kaiser MW, Lyamichev VI, Lyamicheva N;
 XX
 DR WPI; 1998-322748/28.
 XX
 PT Thermostable structure-specific nuclease(s) derived from mutant DNA
 PT polymerase(s) - useful for detecting mutant allele(s) or strains of
 PT microorganisms
 XX
 PS Claim 1; Page 328; 472pp; English.
 CC This is the amino acid sequence of a structure specific nuclease
 CC used in the method of the invention. In this process thermostable
 CC structure-specific nucleases are derived from mutant DNA polymerases,
 CC which can be used for detecting mutant alleles or strains of
 CC microorganisms. The structure-specific nucleases can be used in
 CC mixtures, compositions and kits to treat nucleic acid, e.g. for
 CC detection of wild type and mutant alleles of genes, for detection
 CC and/or identification of strains of microorganisms such as bacteria,
 CC fungi, protozoa, especially for detection of RNA viruses such as the

CC hepatitis C virus (HCV).
 SQ Sequence 325 AA;
 Query Match 3.2%; Score 12; DB 19; Length 325;
 Best Local Similarity 100.0%; Pred. No. 0.0042;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 77 EAGIKPVYVFDG 88
 |||||||||
 Db 69 EAGIKPVYVFDG 80
 RESULT 11
 AAM59953
 ID AAM59953 standard; Protein; 326 AA.
 AC AAM59953;
 XX
 DT 22-DEC-1998 (first entry)
 XX
 DE Amino acid sequence of the structure specific nuclease 5.
 XX
 KW Tag mutant gene; thermostable; nuclease; mutant; DNA polymerase;
 KM bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
 XX
 OS Synthetic.
 XX
 PN WO9823774-A1.
 XX
 PD 04-JUN-1998.
 XX
 PF 26-NOV-1997; 97WO-US21783.
 XX
 PR 02-DEC-1996; 96US-0758314.
 PR 29-NOV-1996; 96US-0757653.
 XX
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Kaiser MW, Lyamichev VI, Lyamicheva N;
 XX
 DR WPI; 1998-322748/28.
 XX
 PT Thermostable structure-specific nuclease(s) derived from mutant DNA
 PT polymerase(s) - useful for detecting mutant allele(s) or strains of
 PT microorganisms
 XX
 PS Claim 1; Pages 331-332; 472pp; English.
 CC This is the amino acid sequence of a structure specific nuclease
 CC used in the method of the invention. In this process thermostable
 CC structure-specific nucleases are derived from mutant DNA polymerases,
 CC which can be used for detecting mutant alleles or strains of
 CC microorganisms. The structure-specific nucleases can be used in
 CC mixtures, compositions and kits to treat nucleic acid, e.g. for
 CC detection of wild type and mutant alleles of genes, for detection
 CC and/or identification of strains of microorganisms such as bacteria,
 CC fungi, protozoa, especially for detection of RNA viruses such as the
 CC hepatitis C virus (HCV).
 XX
 SQ Sequence 326 AA;
 Query Match 3.2%; Score 12; DB 19; Length 326;
 Best Local Similarity 100.0%; Pred. No. 0.0042;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 77 EAGIKPVYVFDG 88
 |||||||||
 Db 70 EAGIKPVYVFDG 81
 RESULT 12
 AAM59951


```

ID  AAW59951 standard; Protein; 332 AA.
XX
XX  AAW59951;
XX
XX  22-DEC-1998 (first entry)
XX
XX  Amino acid sequence of the structure specific nuclease 3.
XX
XX  Tag mutant gene: thermostable; nuclease; mutant; DNA polymerase;
XX  bacteria; protozoa; fungi; RNA virus; hepatitis C virus; HCV.
XX
XX  Synthetic.
XX
XX  MO9823774-A1.
XX
XX  04-JUN-1998.
XX
XX  26-NOV-1997; 97WO-US21783.
XX
XX  02-DEC-1996; 96US-0758314.
XX  29-NOV-1996; 96US-0757653.
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX  Kaiser MW, Lyamichev VI, Lyamicheva N;
XX
XX  WPI: 1998-322748/28.
XX
XX  Thermostable structure-specific nuclease(s) derived from mutant DNA
XX  polymerase(s) - useful for detecting mutant allele(s) or strains of
XX  microorganisms
XX
XX  Claim 1; Pages 329-330; 472pp; English.
XX
XX  This is the amino acid sequence of a structure specific nuclease
XX  used in the method of the invention. In this process thermostable
XX  structure-specific nucleases are derived from mutant DNA polymerases,
XX  which can be used for detecting mutant alleles or strains of
XX  microorganisms. The structure-specific nucleases can be used in
XX  mixtures, compositions and kits to treat nucleic acid, e.g. for
XX  detection of wild type and mutant alleles of genes, for detection
XX  and/or identification of strains of microorganisms such as bacteria,
XX  fungi, protozoa, especially for detection of RNA viruses such as the
XX  hepatitis C virus (HCV).
XX
XX  Sequence 332 AA:
XX
XX  Query Match 3.2%; Score 12; DB 19; Length 332;
XX  Best Local Similarity 100.0%; Pred. No. 0.0043;
XX  Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 77 EAGIKRVYVFDG 88
DB 70 EAGIKRVYVFDG 81

```

```

PD 31-JUL-1997.
XX
XX  22-JAN-1997; 97WO-US01072.
XX
XX  02-DEC-1996; 96US-0759038.
XX  24-JAN-1996; 96US-0599491.
XX  12-JUL-1996; 96US-0682853.
XX  29-NOV-1996; 96US-0756386.
XX  02-DEC-1996; 96US-0758314.
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX  Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI;
XX  Olive DM, Prudent JR;
XX
XX  WPI: 1997-393613/36.
XX  N-PSDB; AAT76685.
XX
XX  Thermostable structure-specific nuclease(s) - used for detection and
XX  characterisation of nucleic acid sequences and variations in nucleic
XX  acid sequences
XX
XX  Example 28; Page 283-285; 457pp; English.
XX
XX  This sequence comprises Pyrococcus furiosus (Pfu) FEN-1
XX  endonuclease. Large-scale production of the enzyme was performed
XX  using E. coli host cells transformed with a vector carrying the Pfu
XX  FEN-1 coding sequence (see AAT76685). Pfu FEN-1 is a thermostable
XX  enzyme. It can be used in novel methods for the detection and
XX  characterisation of nucleic acid sequences and variations in
XX  nucleic acid sequences.
XX
XX  Sequence 340 AA:
XX
XX  Query Match 3.2%; Score 12; DB 18; Length 340;
XX  Best Local Similarity 100.0%; Pred. No. 0.0044;
XX  Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 77 EAGIKRVYVFDG 88
DB 70 EAGIKRVYVFDG 81

```

```

RESULT 13
AAW24216
ID  AAW24216 standard; Protein; 340 AA.
XX
XX  AAW24216;
XX
XX  14-APR-1998 (first entry)
XX
XX  Pyrococcus furiosus FEN-1 endonuclease.
XX
XX  Nucleic acid cleavage; 5' nuclease; DNA cleavage; RNA cleavage;
XX  invader directed cleavage; FEN-1; endonuclease.
XX
XX  Pyrococcus furiosus.
XX
XX  WO9727214-A1.
XX

```

```

RESULT 14
AAW79970
ID  AAW79970 standard; Protein; 340 AA.
XX
XX  AAW79970;
XX
XX  02-FEB-1999 (first entry)
XX
XX  Pyrococcus furiosus FEN-1 endonuclease.
XX
XX  Nucleic acid detection; multiple sequential invasive cleavage;
XX  FEN-1; endonuclease; nuclease.
XX
XX  Pyrococcus furiosus.
XX
XX  WO9842873-A1.
XX
XX  01-OCT-1998.
XX
XX  24-MAR-1998; 98WO-US05809.
XX
XX  24-MAR-1997; 97US-0823516.
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX  Brow MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;
XX  Mast AL, Vavra SH;
XX
XX  WPI: 1998-557036/47.
XX  N-PSDB; AAV65840.
XX

```

XX Detecting target nucleic acid by sequence-specific cleavage of
PT complex with two specific oligonucleotides - used to detect
PT cytomegalovirus DNA
XX
PS Example 28b, Page 316-317; 524pp; English.
XX
CC This is the amino acid sequence of FEN-1 endonuclease of
CC *Pyrococcus furiosus* (Pfu). FEN-1 DNA (see AAV65840) has been
CC ligated into vector for expression in *E. coli* cells. The invention
CC relates to means for the detection and characterization of nucleic
CC acid sequences, and variations in nucleic acid cleavage structure on
CC relates to methods for forming a nucleic acid cleavage structure on
CC a target sequence and cleaving this structure in a site-specific
CC manner, preferably using a thermostable structure-specific nuclease
CC such as FEN-1. Cleavage of the cleavage structure by the
CC nuclease indicates the presence of specific nucleic acid sequences
CC or specific variants. The invention further relates to methods for
CC the separation of nucleic acid molecules based on charge, methods
CC for the detection of non-target cleavage products via the formation
CC of a complete and activated protein binding region, and methods for
CC the detection of nucleic acid from various viruses (e.g. human
CC cytomegalovirus) in a sample. The method amplifies the detection
CC molecule rather than the target itself, is less subject to
CC contamination than exponential amplification processes, and allows
CC many targets to be analysed in a single reaction.
CC
SQ Sequence 340 AA:

Query Match 3.2%; Score 12; DB 19; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPYVFDG 88
|||||
Db 70 EAGIKPYVFDG 81

RESULT 15
AAW59940
ID AAW59940 standard; Protein; 340 AA.
XX
AC AAW59940;
XX
DT 21-DEC-1998 (first entry)
XX
DE Amino acid sequence of the Pfu FEN-1 endonuclease.
XX
KW Mja FEN-1 endonuclease; Tag gene; structure-specific nuclease;
KW mutant; DNA polymerase; bacteria; fungi; protozoa; RNA virus;
KW hepatitis C virus; HCV; thermostable.
XX
OS *Pyrococcus furiosus*.
XX
PN WO9823774-A1.
XX
PD 04-JUN-1998.
XX
PF 26-NOV-1997; 97WO-US21783.
XX
PR 02-DEC-1996; 96US-0758314.
PR 29-NOV-1996; 96US-0757653.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Kaiser MW, Lyamichev VI, Lyamicheva N;
XX
DR WPI, 1998-322748/28.
XX
DR N-PSDB; AAV53951.
XX
PT Thermostable structure-specific nuclease(s) derived from mutant DNA
PT polymerase(s) - useful for detecting mutant allele(s) or strains of
PT microorganisms

XX
PS Example 29; Pages 280-281; 472pp; English.
XX
CC This is the amino acid sequence encoding the Pfu FEN-1 endonuclease,
CC used in the method of the invention. In this process thermostable
CC structure-specific nucleases are derived from mutant DNA polymerases,
CC which can be used for detecting mutant alleles or strains of
CC microorganisms. The structure-specific nucleases can be used in
CC mixtures, compositions and kits to treat nucleic acid, e.g. for
CC detection of wild type and mutant alleles of genes, for detection
CC and/or identification of strains of microorganisms, such as bacteria,
CC fungi, protozoa, especially for detection of RNA viruses such as the
CC hepatitis C virus (HCV).
XX
SQ Sequence 340 AA:

Query Match 3.2%; Score 12; DB 19; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPYVFDG 88
|||||
Db 70 EAGIKPYVFDG 81

RESULT 16
AAW92511
ID AAW92511 standard; Protein; 16 AA.
XX
AC AAW92511;
XX
DT 23-APR-1999 (first entry)
XX
DE Human FEN-1 protein epitope #2.
XX
KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
KW neoplasma; antineoplastic agent; cleavage; epitope.
XX
OS Homo sapiens.
XX
PN US5874283-A.
XX
PD 23-FEB-1999.
XX
PF 30-MAY-1995; 95US-0455968.
XX
PR 30-MAY-1995; 95US-0455968.
XX
PA (HARR/) HARRINGTON J J.
PA (HSIE/) HSIEH C.
PA (LIEB/) LIEBER M R.
XX
PI Harrington JJ, Hsieh C, Lieber MR;
XX
DR WPI; 1999-179985/15.
XX
PT DNA encoding flap endonuclease polypeptides - useful for producing
PT e.g. recombinant polypeptides
XX
PS Disclosure; Column 22; 58pp; English.
XX
CC AAW92410-W92522 represent epitopes from the human FEN-1 (flap
CC endonuclease) protein. This protein can be used in methods for detecting
CC a pathological condition in a patient, for diagnostic purposes, for
CC screening for antineoplastic agents and carcinogens, for diagnostic
CC staging of neoplasia, for producing recombinant flap endonuclease for use
CC as research or diagnostic reagents, for producing antibodies reactive
CC with the novel polypeptides, for producing transgenic nonhuman animals
CC expressing the novel polypeptides encoded by a transgene. The invention
CC also provides novel molecular cloning techniques and reagents involving
CC cleavage of a flap or nick with a flap endonuclease.
XX
SQ Sequence 16 AA:

Query Match 2.4%; Score 9; DB 20; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 SMSIIOFLI 44
 |||||||||
 DB 1 SMSIYOFLI 9

RESULT 17

AAW92506
 ID AAW92506 standard; Protein; 382 AA.

XX AAW92506;

DT 23-APR-1999 (first entry)

DE Yeast FEN-1 protein.

XX FEN-1; Yeast; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasma; antineoplastic agent; cleavage.

OS Saccharomycetes cerevisiae.

XX US5874283-A.

PD 23-FEB-1999.

PF 30-MAY-1995; 95US-0455968.

PR 30-MAY-1995; 95US-0455968.

XX (HARR/) HARRINGTON J J.

PA (HSTE/) HSTEH C.
 (LIEB/) LIEBER M R.

XX Harrington JJ, Hsieh C, Lieber MR;

DR WPI: 1999-179985/15.

DR N-PSDB; AAK02109.

PT DNA encoding flap endonuclease polypeptides - useful for producing

PS e.g. recombinant polypeptides
 Disclosure: Fig 3a; 58pp; English.

XX This sequence represents a yeast FEN-1 (flap endonuclease) protein. This
 CC protein is used in a method to isolate novel human FEN-1 proteins for
 CC detecting a pathological condition in a patient, for diagnostic purposes,
 CC for screening for antineoplastic agents and carcinogens, for diagnostic
 CC staging of neoplasia, for producing recombinant flap endonuclease for use
 CC as research or diagnostic reagents, for producing antibodies reactive
 CC with the novel polypeptides, for producing transgenic nonhuman animals
 CC expressing the novel polypeptides encoded by a transgene. The invention
 CC also provides novel molecular cloning techniques and reagents involving
 CC cleavage of a flap or nick with a flap endonuclease.

XX SQ Sequence 382 AA;

Query Match 2.4%; Score 9; DB 20; Length 382;

Best Local Similarity 100.0%; Pred. No. 4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 84 YVFDGKPPD 92
 |||||||||
 DB 84 YVFDGKPPD 92

RESULT 18

AAK10723

ID AAK10723 standard; Protein; 145 AA.
 XX

AC AAK10723;

DT 22-APR-1991 (first entry)

DE (Lys79)IL-3.

KW Interleukin-3; mutant; antagonist; cytopenia; immunodeficiency;

XX immunosuppression; charge reversal.

XX Key location/Qualifiers

FT Peptide 1..12
 FT /label= signal sequence

FT Mutation 91
 FT /label= Lys -> Glu

XX EP413383-A.

PD 20-FEB-1991.

PF 02-AUG-1990; 90EP-0202117.

PR 14-SEP-1989; 89EP-0202331.

PR 14-AUG-1989; 89EP-0202082.

XX 02-AUG-1990; 90EP-0202117.

PA (KONN) GIST-BROCADES NV.

PI Dorsers LCI, Vanleen RW;

DR WPI: 1991-052685/08.

XX New interleukin-3 mutants - having deletions covering complete

PT coding sequence while retaining biological activity, have

XX antagonistic effect.

XX Claim 6; Page 18; 26pp; English.

XX The mutant was prep'd. by recombinant DNA techniques using mutagenic
 CC primers to change the codons of cDNA encoding IL-3. The peptide
 CC pref. also has a deletion at the N-terminus (AAs 1-14) and/or the
 CC C-terminus (AAs 116-133,120-130 and 130-133). The peptide has
 CC antagonistic activity and can be used for therapeutic and diagnostic
 CC purposes.
 CC See also AAK10711-R10731 and AAK11051-R11055.

XX SQ Sequence 145 AA;

Query Match 2.1%; Score 8; DB 12; Length 145;

Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 257 IESILENL 264
 |||||||||
 DB 86 IESILENL 93

RESULT 19

ABG08130
 ID ABG08130 standard; Protein; 338 AA.

XX AC ABG08130;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8121.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.
 XX

PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-0508631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0549167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS72317.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 38489; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 338 AA:
 XX
 Query Match 2.1%; Score 8; DB 22; Length 338;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 123 AIEKLSKR 130
 |||||||
 Db 273 AIEKLSKR 280
 RESULT 20
 AAY03778
 ID AAY03778 standard; Protein: 343 AA.
 XX
 AC AAY03778;
 XX
 DT 14-JUN-1999 (first entry)
 XX
 DE Pyrococcus heat resistant Flap endonuclease.
 XX
 KM Flap endonuclease; thermophile; thermostable; heat resistant; enzyme;
 KM genetic engineering; recombination; gene shuffling.
 XX
 OS Pyrococcus sp.
 XX
 PN JP11075849-A.
 XX
 PD 23-MAR-1999.

PF 04-SEP-1997; 97JP-0239440.
 XX
 PR 04-SEP-1997; 97JP-0239440.
 XX
 PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
 XX
 DR WPI: 1999-257697/22.
 DR N-PSDB; AAX31850.
 XX
 PT Heat resistant Flap endonuclease - derived from an extreme
 PT thermophile of Pyrococcus genus
 XX
 PS Claim 2; Page 5-6; 8pp; Japanese.
 XX
 CC This represents a heat resistant Flap endonuclease derived from an
 CC extreme thermophile of Pyrococcus genus. Transformants containing a
 CC vector comprising the DNA are cultured for the recombinant expression of
 CC the thermostable Flap endonuclease. The enzyme is thermostable (having
 CC an optimum temperature of at least of 75 deg C) and can be used in
 CC standard genetic engineering techniques, especially for the development
 CC of new methods for efficiently carrying out homologous recombination and
 CC gene shuffling.
 XX
 SQ Sequence 343 AA:
 XX
 Query Match 2.1%; Score 8; DB 20; Length 343;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 84 YVFDGKRP 91
 |||||||
 Db 77 YVFDGKRP 84
 RESULT 21
 AAP70501
 ID AAP70501 standard; Protein: 373 AA.
 XX
 AC AAP70501;
 XX
 DT 13-MAY-1991 (first entry)
 XX
 DE Chinese hamster glutamine synthetase gene product.
 XX
 KW DHFR/MTX.
 XX
 OS Cricetus griseus.
 XX
 PN W08704462-A.
 XX
 PD 30-JUL-1987.
 XX
 PF 23-JAN-1987; 87WO-G000039.
 XX
 PR 23-JAN-1986; 86GB-0001597.
 XX
 PA (CELL-) CELTECH LTD.
 PA (UNIU) UNIV OF GLASGOW.
 PA (WILS/) WILSON RH.
 XX
 PI Wilson RH, Bebbington CR;
 XX
 DR WPI: 1987-221263/31.
 DR N-PSDB; AAN70819.
 XX
 PT Recombinant DNA which encodes glutamine synthetase - used esp. in
 PT co-amplification of non-selected genes and in transforming host
 PT cell lines.
 XX
 PS Claim 6; Fig 2a-e; 54pp; English.
 XX
 CC Sequence encoding the product may be used as a hybridisation probe in
 CC the isolation of the GS gene in other species, and as a diagnostic tool

CC in the detection of diseases altering GS expression. A second use is
CC as a selectable marker in recombinant DNA technology eg. co-amplification
CC selection, or transformation to glutamine independence.

XX
SQ Sequence 373 AA;

Query Match 2.1%; Score 8; DB 8; Length 373;

Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 AIEKLSKR 130
|||||

DB 273 AIEKLSKR 280

RESULT 22

ABB59358 ID ABB59358 standard; Protein; 399 AA.

XX ABB59358;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 4866.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL03461.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Disclosure; SEQ ID NO 4866; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB101840-ABL16175), expressed DNA

XX sequences (AB101840-ABL16175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

RESULT 23

ABB65740 ID ABB65740 standard; Protein; 399 AA.

XX ABB65740;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 24012.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL09843.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Disclosure; SEQ ID NO 24012; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB101840-ABL16175), expressed DNA

XX sequences (AB101840-ABL16175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 399 AA;

XX Query Match 2.1%; Score 8; DB 22; Length 399;

XX Best Local Similarity 100.0%; Pred. No. 39;

XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 AIEKLSKR 130
|||||

DB 309 AIEKLSKR 316

RESULT 24

AAV95714 ID AAV95714 standard; Protein; 515 AA.

XX AAV95714;

XX 25-OCT-2000 (first entry)

XX Cosmid CHRIM5 encoded protein P20-9r.

XX Cosmid CHRIM5; nematocidae; nematode; biological control agent;

```

KW transgenic plant; helminthiasis; P20-9r.
XX
OS Xenorhabdus bovienii.
XX
PN WO200042855-A1.
XX
PD 27-JUL-2000.
XX
PF 24-JAN-2000; 2000MO-GB00219.
XX
PR 22-JAN-1999; 99GB-0001499.
XX
PA (HORT-) HORTICULTURE RES INT.
XX
PI Morgan JAW, Jarrett P, Ellis D, Ousley MA.
XX
DR WPI; 2000-499157/44.
XX
DR N-PSDB; AAA50029.
XX
PT Novel composition used to control parasitic nematodes, especially in
PT plants such as maize, cotton, soya, and rice, comprises a bacterium
PT which is a symbiont of an entomopathogenic nematode -
XX
PS Example 6; Page 44; 74pp; English.
XX
CC The present sequence is that of protein P20-9r encoded by an open
CC reading frame identified in cosmid CHRIM5 (see AAA50029). CHRIM5 was
CC obtained by ligating Xenorhabdus bovienii strain 173 (NCIMB 40986)
CC Sau3A-digested DNA fragments into the BamHI site of the Stratagene
CC cosmid vector Supercos1, packaging into Escherichia coli XL Blue 1,
CC and screening for nematocidal activity against Caenorhabditis elegans.
CC Analysis of the DNA indicated a number of open reading frames for
CC CC AAAY5685-Y95735). Nematodes can be controlled through the use of
CC bacteria associated symbiotically with an entomopathogenic nematode.
CC Such bacteria include Xenorhabdus and Photorhabdus spp. such as X.
CC bovienii strain 173. The symbiont bacteria, an engineered
CC bacterium, or a nematocidal protein obtained from such bacteria,
CC can be used to control helminthiasis in a human or domesticated
CC animal or for the control of plant pathogen nematodes. Also
CC claimed are vectors for expressing nematocidal proteins in host
CC cells, and transgenic plants.
CC
SQ Sequence 515 AA;

```

Query Match 2.1%; Score 8; DB 21; Length 515;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 144 RLRLMGV 151
   |||||
DQ 480 RLRLMGV 487

```

RESULT 25
AAV58997
ID AAV58997 standard; Protein: 687 AA.
XX
AC AAV58997;
XX
DT 23-MAY-2000 (first entry)
XX
DE Soybean sulfite reductase.
XX
KM Soybean; sulfite reductase; sulfate assimilation; transgenic plant.
XX
OS Glycine max.
XX
PN WO200004162-A2.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15812.

```

XX 14-JUL-1998; 98US-0092833.
PR
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
PI Falco SC, Allen SM;
XX
DR WPI; 2000-182428/16.
XX
DR N-PSDB; AAZ58541.
XX
PT Nucleic acid fragments encoding sulfate assimilation proteins in plants
PT and seeds useful as probes for isolating cDNAs and genes encoding
PT homologous proteins, in producing transgenic plants -
XX
PS Claim 6; Page 31-33; 40pp; English.
XX
CC The present sequence is that of soybean sulfite reductase, as
CC deduced from an isolated cDNA clone (see AAZ58541). The invention
CC relates to isolated nucleic acid fragments (see AAZ58540-42) encoding
CC sulfite reductase of rice, soybean and wheat (see AAY38996-98). It
CC also relates to the construction of a chimeric gene encoding all or
CC a portion of the sulfite reductase in sense or antisense orientation,
CC where expression of the chimeric gene results in production of altered
CC levels of the sulfate assimilation protein in a transformed host cell.
CC The invention also includes transformed plants, and seeds derived from
CC them. Manipulation of enzymes involved in sulfate assimilation may
CC increase levels of S-containing compounds, including cysteine and
CC methionine, for use in animal feeds.
CC
SQ Sequence 687 AA;

```

Query Match 2.1%; Score 8; DB 21; Length 687;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 348 ATTSAPLK 355
   |||||
DQ 8 ATTSAPLK 15

```

RESULT 26
AAV43478
ID AAV43478 standard; Protein: 770 AA.
XX
AC AAV43478;
XX
DT 26-JAN-2000 (first entry)
XX
DE Amino acid sequence of a SC20 protein of soybean.
XX
KM Promoter; seed coat specificity; SC20 protein;
KM herbicide resistance gene; viral coat protein; biological control;
KM Bt toxin; seed taste.
XX
OS Glycine max.
XX
FH Key
FH Region
FT 4..7
FT /note= "consensus sequence for N-glycosylation"
FT 30..33
FT /note= "consensus sequence for N-glycosylation"
FT 188..191
FT /note= "consensus sequence for N-glycosylation"
FT 388..394
FT /note= "consensus sequence for N-glycosylation"
FT 475..478
FT /note= "consensus sequence for N-glycosylation"
FT 533..536
FT /note= "consensus sequence for N-glycosylation"
FT 644..647
FT /note= "consensus sequence for N-glycosylation"
FT 681..684
FT /note= "consensus sequence for N-glycosylation"

FT Region 736..739
 FT /note= "consensus sequence for N-glycosylation"
 XX
 PN W09953067-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 13-APR-1999; 99WO-CA00293.
 XX
 PR 13-APR-1998; 98US-0059090.
 XX
 PA (MTAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 XX
 PI Mixl B, Gijzen M, Miller S, Bowman L, Batchelor A, Hu M;
 PI Bouillier K;
 DR WPI: 1999-611304/52.
 DR N-PSDB: AA229905.
 XX
 PT Novel promoter sequences and genes useful for inducing expression of
 PT genes in plant seed coats -
 XX
 PS Example: Fig 23(B); 155pp; English.
 XX
 CC The present sequence represents a SC20 protein. The promoter of
 CC this gene is differentially expressed in seed coat tissues,
 CC specifically within the outer integument, the inner integument, the
 CC thick walled parenchyma, the thin walled parenchyma, the endodermis,
 CC the hourglass cells, the palisade, the stellate parenchyma, or the
 CC membranous endocarp associated with the seed coat. The seed-coat
 CC promoters may be used to express proteins of interest in seed coat
 CC tissues. Genes of interest include but are not restricted to herbicide
 CC resistance genes, genes encoding viral coat proteins, or genes encoding
 CC proteins conferring biological control of pests or pathogens, e.g. a Bt
 CC toxin. Other genes that may be expressed include proteins that alter the
 CC taste of the seed and/or affect the nutritive value of the seed.
 CC
 SQ Sequence 770 AA:
 XX
 Query Match 2.1%; Score 8; DB 20; Length 770;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 152 PYVEAPSE 159
 |||||
 DB 707 PYVEAPSE 714
 .
 RESULT 27
 AAM96822
 ID AAM96822 standard; Peptide; 14 AA.
 XX
 AC AAM96822;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Human peptide #97 encoded by a SNP oligonucleotide.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease.
 XX
 OS Homo sapiens.
 XX
 PN W0200147944-A2.
 XX
 PD 05-JUL-2001.
 XX

PF 28-DEC-2000; 2000WO-US35498.
 XX
 XX 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 DR WPI: 2001-465210/50.
 XX
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 XX
 PS Disclosure: Page 3689; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesins, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 CC
 SQ Sequence 14 AA:
 XX
 Query Match 1.8%; Score 7; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 179 EDMDSLTL 185
 |||||
 DB 4 EDMDSLTL 10
 .
 RESULT 28
 AAM92515
 ID AAM92515 standard; Protein; 25 AA.
 XX
 AC AAM92515;
 XX
 DT 23-APR-1999 (first entry)
 XX
 DE Human FEN-1 protein epitope #6.
 XX
 KW FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasia; antineoplastic agent; cleavage; epitope.
 XX
 OS Homo sapiens.
 XX
 PN US5874283-A.
 XX
 PD 23-FEB-1999.
 XX
 PF 30-MAY-1995; 95US-0455968.
 XX
 PR 30-MAY-1995; 95US-0455968.
 XX
 PA (HARR/) HARRINGTON J J.
 PA (HSIEH/) HSIEH C.
 PA (LIEBER/) LIEBER M R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX

DR	WP1: 1999-179985/15.
XX	
PX	DNA encoding flap endonuclease polypeptides - useful for producing
PT	e.g. recombinant polypeptides
XX	
PS	Disclosure; Column 22; 58pp; English.
XX	
CC	AAM92410-w92552 represent epitopes from the human FEN-1 (flap
CC	endonuclease) protein. This protein can be used in methods for detecting
CC	a pathological condition in a patient, for diagnostic purposes, for
CC	screening for antineoplastic agents and carcinogens, for diagnostic
CC	staging of neoplasms, for producing recombinant flap endonuclease for use
CC	as research or diagnostic reagents, for producing antibodies reactive
CC	with the novel polypeptides, for producing transgenic nonhuman animals
CC	expressing the novel polypeptides encoded by a transgene. The invention
CC	also provides novel molecular cloning techniques and reagents involving
CC	cleavage of a flap or nick with a flap endonuclease.
SQ	Sequence 25 AA;
Query Match	1.8%; Score 7; DB 20; Length 25;
Best Local Similarity	100.0%; Pred. No. 28;
Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	156 APSEAEA 162
Db	12 APSEAEA 18
RESULT 29	
ID	AAM90755 standard; Protein; 49 AA.
XX	
AC	AAM90755;
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen SEQ ID NO:18348.
XX	
KW	Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis.
XX	
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	
PD	09-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01354.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198125.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	26-JUL-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	

[illegible]

PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0251719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI: 2001-483426/52.
 DR N-PSDB; AAK63536.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Claim 11: SEQ ID NO 18348: 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX

SQL Sequence 49 AA:
 Query Match 1.8%; Score 7; DB 22; Length 49;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 346 PTAITSA 352
 |||||
 Db 31 PTAITSA 37
 RESULT 30
 AAB34731
 ID AAB34731 standard; Protein; 61 AA.
 AC AAB34731;
 XX
 DT 26-JAN-2001 (first entry)
 DE
 XX Human secreted protein encoded by DNA clone vq13 1.
 DE
 XX Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;
 KW systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;
 KW haematopoiesis regulation; tissue growth; wound healing; haemophilia;
 KW Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;
 KW contraceptive; infection; growth inhibition; hyperproliferative disorder;
 KW poriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200055375-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 17-MAR-2000; 2000WO-US07285.
 XX
 PR 17-MAR-1999; 99US-0124808.
 PR 17-MAR-1999; 99US-0124916.
 PR 17-AUG-1999; 99US-0149639.
 PR 01-OCT-1999; 99US-0157247.
 PR 29-NOV-1999; 99US-0167824.
 PR 15-FEB-2000; 2000US-0182711.
 XX
 PA (ALPH-) ALPHAGENE INC.
 XX
 PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 DR WPI: 2000-638211/61.
 DR N-PSDB; AAC59832.
 XX
 PT Novel proteins and polypeptides useful for the treatment of e.g
 PT multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
 PT cancer, Alzheimer's disease, Parkinson's disease, stroke, anemia and
 PT ulcers -
 XX
 PS Claim 98: Page 444-445; 493pp; English.
 XX
 CC This invention relates to 59 human secreted proteins and the nucleotide
 CC sequences encoding them. Sequences AAC59788-C59846 and AAB34687-834745
 CC represent the proteins and their encoding nucleotide sequences, and
 CC sequences AAB34746-B34771 represent fragments of the proteins. Probes
 CC for the DNA sequences are represented by sequences AAC59847-C599596. The
 CC proteins exhibit neuroprotective, dermatological, immunosuppressive,
 CC antiinflammatory, antiaemic, nootropic, antiparkinsonian,
 CC cerebroprotective, haemostatic, vulnerary, cytoprotic, antipruritic,
 CC antibacterial, virucide, and fungicide activity. The proteins and
 CC nucleotide sequences are useful as nutritional sources or supplements
 CC and in research. The proteins are useful for treating immune deficiency
 CC and disorders, which may be genetic or resulting from infections,
 CC autoimmune disorders such as multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, and for treating myeloid or lymphoid
 CC cell deficiencies such as anaemias by regulating haematopoiesis. The
 CC proteins are also useful in compositions for bone, cartilage, tendon,
 CC

CC ligament and/or nerve tissue growth or regeneration, for wound healing,
CC tissue repair and replacement and in the treatment of wounds, incisions
CC and ulcers. Other uses include in the treatment of central and
CC peripheral nervous system and neuropathies, such as Alzheimer's and
CC Parkinson's diseases and Shy-Drager syndrome, and mechanical and
CC traumatic disorders, such as spinal cord disorders, head trauma and
CC stroke. The proteins may also be used as a contraceptive, and for
CC treating coagulation disorders such as haemophilias. The protein and
CC nucleotide sequences with cadherin activity are useful for treating
CC cancer. Other uses for the protein include for inhibiting the growth,
CC infection or function of, or killing, infectious agents such as bacteria,
CC virus, fungi and other parasites, for effecting bodily characteristics
CC such as height, weight, hair colour, effecting biorhythms or cardiac
CC cycles or rhythms, effecting metabolism, catabolism, anabolism,
CC processing, utilization, storage or elimination of dietary fat, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors, effecting
CC behavioural characteristics, providing analgesic effects and for treating
CC hyperproliferative disorders such as psoriasis.

CC Sequence 61 AA;

Query Match 1.8%; Score 7; DB 21; Length 61;
Best Local Similarity 100.0%; Pred. No. 64;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 IEKLSKR 130
| | | | | | |
DB 32 IEKLSKR 38

RESULT 31
AAU61110
ID AAU61110 standard; Protein: 62 AA.

AC AAU61110;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #22006.

XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperkerosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

PN 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Sheiky YAM, Persing DH, Mitcham JL, Wang SS, Bhattacharya A;

XX L'walsonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59613.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX treating acne vulgaris -

CC Example 1; SEQ ID NO 22305; 1069pp; English.
CC Sequences AAU3105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperkerosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 62 AA;

Query Match 1.8%; Score 7; DB 22; Length 62;
Best Local Similarity 100.0%; Pred. No. 65;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 PDEEGLI 307
| | | | | | |
DB 26 PDEEGLI 32

RESULT 32
AAU02011
ID AAU02011 standard; Protein: 72 AA.

AC AAU02011;

DT 29-AUG-2001 (first entry)

DE Human secreted protein encoded by gene #37.

XX
XX Human secreted protein; autoimmune disorder; hyperproliferative disorder;
XX cardiovascular disorder; cerebrovascular disorder; angiogenesis;
XX nervous system disorder; bacterial infection; viral infection;
XX fungal infection; ocular disorder; wound healing; tissue regeneration;
XX epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.

XX Homo sapiens.

XX WO200123598-A1.

PN 05-APR-2001.

PF 26-SEP-2000; 2000WO-US26324.

PR 27-SEP-1999; 99US-0155807.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Komatsoulis G, Ruben SM, Rosen CA;

XX WPI; 2001-281684/29.

XX N-PSDB; AAS03918.

XX Forty one nucleic acid molecules encoding human secreted proteins, useful

XX in the prevention, treatment and diagnosis of cancer, immune disorders,

XX cardiovascular disorders and neurological diseases -

XX Disclosure; Page 512; 518pp; English.

XX The sequence represents a human secreted protein encoded by a nucleic

XX acid of the invention. Secreted proteins and their related nucleic acids
XX can be used in the diagnosis of or susceptibility to a pathological
XX condition by determining the presence or absence of a mutation in a

CC nucleic acid or the presence or amount of expression of a secreted
CC protein. The sequences are used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. The antibodies to the polypeptides can also be used in
CC alleviating symptoms associated with disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent
CC assays (ELISA). The disorders include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The peptides can also
CC be used to aid wound healing and epithelial cell proliferation, to help
CC prevent skin ageing due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues, in chemotaxis and as a food
CC additive or preservative to alter storage capabilities.
XX
SQ Sequence 72 AA;

Query Match 1.8%; Score 7; DB 22; Length 72;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 276 PYOEARR 282
| | | | | | | |
Db 43 PYOEARR 49

RESULT 33
AAG05446
ID AAG05446 standard; Protein: 73 AA.
XX
XX AAG05446;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1855.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159588.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 73;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 KDAIEKL 127
Db 65 KDAIEKL 71

RESULT 34
AAG07581

ID AAG07581 standard; Protein: 73 AA.

XX AC AAG07581;

XX DF 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 4789.

KW Protein identification: signal transduction pathway; metabolic pathway;
KM hydridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP103405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 09-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 06-MAY-1999; 99US-0132487.

XX PR 07-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136621.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141642.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159648.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 73;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127
Db 65 KDAIEKL 71

RESULT 35
AAG08148
ID AAG08148 standard; Protein: 73 AA.

AC AAG08148;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 5572.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147435.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.88; Score 7; DB 21; Length 73;
Best Local Similarity 100.0%; Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 121 KDAIEKL 127
1111111
Db 65. KDAIEKL 71
RESULT .36
AAG08766
ID AAG08766 standard; Protein; 73 AA.
XX
AC AAG08766;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6433.
XX
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147206.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-01593295.
PR 14-OCT-1999; 99US-01593295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 1.0%; Score 7; DB 21; Length 73;
Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127

Db 65 KDAIEKL 71

RESULT 37
AAG36961
ID AAG36961 standard; Protein; 73 AA.
XX AAG36961;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SFO ID NO: 45372.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-030439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 73;
Best Local Similarity 100.0%; Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 KDAIEKL 127
Db 65 KDAIEKL 71

RESULT 38
AAG52931
ID AAG52931 standard; Protein; 73 AA.
XX AAG52931;
AC
XX
DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 67337.
DE
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155489.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 73;
Best Local Similarity 100.0%; Pred No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127
|||
Db 65 KDAIEKL 71

RESULT 39
AAVS9186
ID AAVS9186 standard; peptide; 73 AA.
XX
AC AAVS9186;
XX
DT 28-MAR-2000 (first entry)
XX
DE Alpha-2 antipiasmin (alpha-2AP) peptide mimetic.
XX
KW Blood clot; alpha-2 antipiasmin; alpha-2AP; factor XIIIa; fibrin;
KW lysis; thrombolytic; thrombosis; stroke; myocardial infarction;
XX pulmonary embolism; deep venous thrombosis.
XX

```

OS Homo sapiens.
XX Key Location/Qualifiers
XX Key MISC-difference 6 /label= Trp or Arg
XX FT
XX FT
XX PN W09961072-A2.
XX PD 02-DEC-1999.
XX PE 28-MAY-1999; 99WO-US11843.
XX PR 29-MAY-1998; 98US-0087218.
XX PA (HARD ) HARVARD COLLEGE.
XX PI Reed GL;
XX DR WPI; 2000-116329/10.
XX PT Inhibiting clot formation and promoting therapeutic thrombolysis using
XX PT alpha-2 antipiasmin polypeptides -
XX PS Claim 38; Page 32; 37pp; English.
XX CC The invention relates to a novel method of detecting blood clot formation
XX CC in a mammal that comprises administering a diagnostically effective
XX CC amount of a detectably labeled alpha-2 antipiasmin (alpha-2AP)
XX CC polypeptide. Alpha-2AP polypeptides inhibit factor XIIIa-mediated
XX CC crosslinking of endogenous alpha-2AP with fibrin, to cause thrombi to
XX CC undergo spontaneous physiologic lysis or accelerated lysis when
XX CC administered with therapeutic thrombolytic agents. The alpha-2AP
XX CC polypeptides can be labeled, and used to non-invasively detect the
XX CC formation of clots and for in vivo imaging. The methods may also be used
XX CC to prevent the development of blood clots in patients at risk for stroke,
XX CC myocardial infarction, pulmonary embolism, and deep venous thrombosis.
XX CC The peptide may also be linked to thrombolytic agents and used to
XX CC dissolve blood clots. The present sequence represents a peptide mimetic
XX CC of human alpha-2AP polypeptide.
XX SO
XX Sequence 73 AA;
XX
XX Query Match 1.8%; Score 7; DB 21; Length 73;
XX Best Local Similarity 100.0%; Pred. No. 75;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 243 GGQTALK 249
XX |
XX 31 GGQTALK 37
XX
XX Db
XX
XX RESULT 40
XX AAY59187
XX ID AAY59187 standard; peptide; 73 AA.
XX AC AAY59187;
XX DT 28-MAR-2000 (first entry)
XX DE Alpha-2 antipiasmin (alpha-2AP) peptide mimetic.
XX KW Blood clot; alpha-2 antipiasmin; alpha-2AP; factor XIIIa; fibrin;
XX KW lysis; thrombolytic; thrombosis; stroke; myocardial infarction;
XX KW pulmonary embolism; deep venous thrombosis.
XX OS Homo sapiens.
XX PN W09961072-A2.
XX PD 02-DEC-1999.
XX PF 28-MAY-1999; 99WO-US11843.
XX

```

```

XX 29-MAY-1998; 98US-0087218.
XX (HARD ) HARVARD COLLEGE.
XX PI Reed GL;
XX DR WPI; 2000-116329/10.
XX PT Inhibiting clot formation and promoting therapeutic thrombolysis using
XX PT alpha-2 antipiasmin polypeptides -
XX PS Claim 9; Page 29; 37pp; English.
XX CC The invention relates to a novel method of detecting blood clot formation
XX CC in a mammal that comprises administering a diagnostically effective
XX CC amount of a detectably labeled alpha-2 antipiasmin (alpha-2AP)
XX CC polypeptide. Alpha-2AP polypeptides inhibit factor XIIIa-mediated
XX CC crosslinking of endogenous alpha-2AP with fibrin, to cause thrombi to
XX CC undergo spontaneous physiologic lysis or accelerated lysis when
XX CC administered with therapeutic thrombolytic agents. The alpha-2AP
XX CC polypeptides can be labeled, and used to non-invasively detect the
XX CC formation of clots and for in vivo imaging. The methods may also be used
XX CC to prevent the development of blood clots in patients at risk for
XX CC thrombosis and to treat patient with thrombotic condition such as stroke,
XX CC myocardial infarction, pulmonary embolism, and deep venous thrombosis.
XX CC The peptide may also be linked to thrombolytic agents and used to
XX CC dissolve blood clots. The present sequence represents a peptide mimetic
XX CC of human alpha-2AP polypeptide.
XX SO
XX Sequence 73 AA;
XX
XX Query Match 1.8%; Score 7; DB 21; Length 73;
XX Best Local Similarity 100.0%; Pred. No. 75;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 243 GGQTALK 249
XX |
XX 31 GGQTALK 37
XX
XX Db
XX
XX RESULT 41
XX AAY59188
XX ID AAY59188 standard; peptide; 73 AA.
XX AC AAY59188;
XX DT 28-MAR-2000 (first entry)
XX DE Alpha-2 antipiasmin (alpha-2AP) peptide mimetic.
XX KW Blood clot; alpha-2 antipiasmin; alpha-2AP; factor XIIIa; fibrin;
XX KW lysis; thrombolytic; thrombosis; stroke; myocardial infarction;
XX KW pulmonary embolism; deep venous thrombosis.
XX OS Homo sapiens.
XX PN W09961072-A2.
XX PD 02-DEC-1999.
XX PF 28-MAY-1999; 99WO-US11843.
XX PR 29-MAY-1998; 98US-0087218.
XX PA (HARD ) HARVARD COLLEGE.
XX PI Reed GL;
XX DR WPI; 2000-116329/10.
XX PT Inhibiting clot formation and promoting therapeutic thrombolysis using
XX PT alpha-2 antipiasmin polypeptides -

```

xx Claim 10; Page 29; 37pp; English.
ps
xx
cc The invention relates to a novel method of detecting blood clot formation
cc in a mammal that comprises administering a diagnostically effective
cc amount of a detectably labeled alpha-2 antiplasmin (alpha-2AP)
cc polypeptide. Alpha-2AP polypeptides inhibit factor XIIIa-mediated
cc crosslinking of endogenous alpha-2AP with fibrin, to cause thrombi to
cc undergo spontaneous physiologic lysis or accelerated lysis when
cc administered with therapeutic thrombolytic agents. The alpha-2AP
cc polypeptides can be labeled, and used to non-invasively detect the
cc formation of clots and for in vivo imaging. The methods may also be used
cc to prevent the development of blood clots in patients at risk for
cc thrombosis and to treat patient with thrombotic condition such as stroke,
cc myocardial infarction, pulmonary embolism, and deep venous thrombosis.
cc The peptide may also be linked to thrombolytic agents and used to
cc dissolve blood clots. The present sequence represents a peptide mimetic
cc of human alpha-2AP polypeptide.
cc
xx
SQ Sequence 73 AA:
Query Match 1.8%; Score 7; DB 21; Length 73;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 243 GGQTALK 249
|||||
Db 31 GGQTALK 37
RESULT 42
AAU63518
ID AUA63518 standard; Protein; 75 AA.
xx
AC AUA63518;
xx
DT 27-FEB-2002 (first entry)
xx
DE Propionibacterium acnes immunogenic protein #24414.
xx
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
xx
OS Propionibacterium acnes.
xx
PN WO200181581-A2.
xx
PD 01-NOV-2001.
xx
PF 20-APR-2001; 2001WO-US12865.
xx
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
xx
PA (CORI-) CORIXA CORP.
xx
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A,
PI L'maisonneuve J, Zhang Y, Jen S, Carter D.
xx
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59633.
xx
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
xx
XX Example 1; SEQ ID No 24713; 1069pp; English.
CC Sequences AAU9105-AAU6017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
cc
xx
SQ Sequence 75 AA:
Query Match 1.8%; Score 7; DB 22; Length 75;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 130 RTVKYTR 136
|||||
Db 27 RTVKYTR 33
RESULT 43
AAM84718
ID AAM84718 standard; Protein; 86 AA.
xx
AC AAM84718;
xx
DT 07-NOV-2001 (first entry)
xx
DE Human immune/haematopoietic antigen SEQ ID NO:12311.
xx
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
xx
OS Homo sapiens.
xx
PN WO200157182-A2.
xx
PD 09-AUG-2001.
xx
PF 17-JAN-2001; 2001WO-US01354.
xx
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.

SEQ Sequence 86 AA: 1.8%; Score 7; DB 22; Length 86;
Query Match Best Local Similarity 100.0%; Pred. No. 87;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 LIVGRT 49
|||||
DB 6 LIVGRT 12

RESULT 44
ABG17892
ID ABG17892 standard; Protein; 101 AA.
XX
AC ABG17892;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #17883.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS82079.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20: SEQ ID No 48251; 103pp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 101 AA:

Query Match 1.8%; Score 7; DB 22; Length 101;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 VFAVASE 179
|||||
DB 83 VFAVASE 89

RESULT 45
AAG05445
ID AAG05445 standard; Protein; 106 AA.
XX
AC AAG05445;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1854.
XX
DE Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
OS
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147310.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 7; DB 21; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127
DB 98 KDAIEKL 104

Search completed: November 5, 2002, 16:22:40
Job time : 67 secs
